THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY

CARL MURCHISON

VOLUME 60

1942

Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

TABLE OF CONTENTS

$7.00 per volume
Single numbers $3.50

QUARTERLY
Two volumes per year

March, 1942
Volume 60, First Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

[illegible] be unstamped, this [illegible] the regular [illegible] edition
[illegible] with a designating [illegible] is a special
[illegible] of a bilateral contract, and may not
[illegible] years from date of publication.

Entered as second-class matter, August 1, 1897, at the post-office at
Worcester, Mass., under the Act of March 3, 1879.
Re-entered as second-class matter, May [illegible], at the post-office at
Provincetown, Mass., under the Act of March 3, 1879.

The Journal of Psychology

[illegible]

Carl Murchison

The Journal of Psychology [illegible]

(1) [illegible]

(2) The [illegible] of each [illegible] will [illegible] its [illegible]

(3) [illegible] through the fault of the author, the lag [illegible] than three or four weeks.

(4) [illegible] receive 200 reprints of [illegible] accumulated publications [illegible] of each volume will [illegible] volume. Each half volume [illegible] its stated date.

(5) [illegible] author will be charged the actual manu[illegible] cuts.

(6) [illegible] tion charges will be as follows:

4 page signature	$12
8 " "	[illegible]
[illegible]	[illegible]

These [illegible]

(7) [illegible] January, [illegible]

Two [illegible]

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

Two volumes per year
QUARTERLY

March, 1942
Volume 60, First Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

MARCH, 1942

(Manuscripts are printed in the order of final acceptance)

(OVER)

Provincetown, Massachusetts
Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass., under the Act of March 3, 1879
Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The Journal of Genetic Psychology, 1942, 60, 3-35.

ATTITUDES TOWARD PROPERTY: A GENETIC STUDY BY THE PAIRED-COMPARISONS RATING OF OFFENSES*

Department of Psychology, Northwestern University

JOHN C. EBERHART[1]

A. INTRODUCTION

The cultural conditioning which every neonate goes through in the process of growing up has many aspects. One of the most important of these concerns property and the rights accruing to it. In our society the rules concerning the ownership of property define certain adaptations which must be made if the individual is to avoid an important source of conflict.

The term "property" in the legal sense includes tangible objects, such as real estate and personal possessions, and certain intangible rights, such as patents and franchises. The present study deals only with tangible possessions, and specifically with those which come within the experience of elementary and secondary school boys of metropolitan Chicago.

The category of possessions which are subsumed by the term "property" is so broad that probably no one has an attitude toward "property" in general, in the sense that we can speak of an attitude toward one's mother, toward the New Deal, or toward war. These latter attitudes can be described as existing, in one dimension, at some point on a pro-anti continuum. They can be so conceived of because their objects (mother, New Deal, war) are in the thinking of most people simple, discrete unities, which can be meaningfully liked or disliked, approved or condemned.

There are probably some differences between adults in the degree to which they value property. It is not uncommon to find an adult whose philosophy of life includes a derogation of the value of material possessions for himself although in our culture it is more usual

*Received in the Editorial Office on September 2, 1940.

[1]The writer is indebted to Prof. A. R. Gilliland and to Dr. H. D. Williams, both of whom made helpful suggestions during the course of this study. The financial assistance of the Payne Fund during the early stages is also gratefully acknowledged.

to find that adults set considerable store by the possessions they are able to accumulate. It would, then, be meaningful to work out a scale to measure attitudes toward "the importance of property for a happy life," or some similarly phrased relationship.

Such a question would not be meaningful to most children, nor indeed to many adults. Children are forced to recognize from an early age the existence of something called "property rights" or the "rights of ownership," and they find out to their sorrow that violations of such rights bring unpleasant consequences. The present study is concerned with finding out how the first-grade boy conceives of property rights, and how this conception changes with age and with progress through the school system.

It is unfortunate but true that children do not willingly incriminate themselves. As is the case also with adults, direct questions about moral issues can be expected to produce evasive and distorted answers if the subject feels any social disapproval toward his opinions. The ordinary attitude-scale techniques are therefore not useful in investigating attitudes toward property rights. For this study the rating of offenses against property according to seriousness, by the paired-comparisons method, was chosen as an indirect approach to an understanding of the boy's conception of property rights.

The problems of this study may be summarized as follows:

1. What changes take place in boys' rankings of property offenses from the first grade to the twelfth grade, and how do these rankings compare with those of adults?

2. What factors are responsible for the ranking of offenses against property?

B. Subjects

All subjects were males. Subjects from grades one through twelve were taken from three schools in one neighborhood in Chicago—the Skinner and Otis grammar schools, and the Crane Technical High School. Differences between the three schools in socio-economic status and in amount of delinquency were slight. The area in which they are located is Chicago's near west side, and is one of the "interstitial areas" described by Shaw (25). The delinquency rates of the districts in which these three schools are located are about midway between the highest and the lowest.

Approximately 100 boys were tested in each grade from the fifth through the twelfth. Because boys below the fifth grade had to be

tested individually the number of subjects in grades one and three was lower than in the higher grades, and grades two and four were omitted altogether. Table 1 shows the number of subjects in each grade, and the age distribution of those subjects.

TABLE 1
NUMBERS OF SUBJECTS USED IN EACH GRADE FROM 1 THROUGH 12, DISTRIBUTED ACCORDING TO AGE

	Grades 1	3	5	6	7	8	9	10	11	12	Age totals
Age											
6	43										43
7	23										23
8	4	19									23
9		16	6								22
10		7	22	5							34
11		1	30	21	4						56
12			29	29	28	2	2				90
13			4	28	20	17	16	4			89
14			1	6	21	19	39	22	3		111
15			2	4	5	21	21	37	24	2	116
16				1	3	15	6	23	53	27	128
17					1		4	7	23	42	77
18								3	1	27	31
19										6	6
20											
21								1			1
Grade totals	70*	43	94	94	82	74	88	97	104	104	850

*Rankings of 14 subjects in the first grade were discarded because of imperfections, as explained by the writer in a previous paper (9). Scores for the first grade are thus based on 56 cases.

The adults used were former members of the writer's psychology class in an institute for the unemployed on Chicago's west side. Thirty-two out of 75 replied to a mailed questionnaire. The ages of these subjects ranged from 18 to 30 with a mean at 21, and most were high school graduates.

C. PROCEDURE

1. *The Statement Test*

Twenty offenses were chosen from the life histories of delinquents secured by Clifford R. Shaw and his associates and filed at the Institute for Juvenile Research in Chicago. These 20 offenses are given in the following list:

1. To swipe $1.00 from your brother's bank at home.
2. To help yourself to chocolates from a box in your sister's room.
3. To borrow your brother's baseball without asking.
4. To swipe your mother's wrist watch and pawn it.
5. To lift $1.00 from your father's pants' pocket when taking the pants to the tailor.
6. To keep $1.00 you find on the street without trying to find the owner.
7. To take a wheel from a wagon you find in the alley.
8. To keep a candy package you find after it has fallen from a truck.
9. To keep $1.00 you see a man drop from his pocket.
10. To keep a ball and glove you find in the school yard.
11. To swipe flowers from a park.
12. To snitch three tickets from a movie cashier.
13. To steal candy and cigarettes from a boxcar.
14. To ride on the street car for half fare when you should pay full fare.
15. To sneak by an "L" cashier without paying.
16. To take a wagon from a boy's back yard.
17. To swipe a dollar from your boss's desk.
18. To snitch fruit from a peddler's stand.
19. To swipe and sell lead pipes from an old warehouse.
20. To sneak a rubber ball from a dime store counter.

It will be seen that these offenses vary in the value of the property involved, in the term used to describe the offense, and in the likelihood of punishment. The offenses may further be grouped into four categories of ownership: property in the home (Offenses 1, 2, 3, 4, 5); lost property (Offenses 6, 7, 8, and 10); property having many owners (Offenses 11, 12, 13, 14, 15, 19, 20); and property owned by one person (Offenses 9, 16, 17, 18).

For securing rankings from all subjects except those in grades one and three a mimeographed booklet was devised in which each offense was paired with every other offense. The number of pairs resulting was 190, and the booklet was 8½ pages long. This booklet will hereafter be known as the "statement test" to distinguish it from a "picture test" which was devised for use with grades one and three. The blank is not properly a "test," but the term is convenient.

The arrangement of the offenses in the statement test was made

as follows: (*a*) the number of every offense was written down beside the number of every other offense, each pair of numbers on a separate slip of paper; (*b*) each offense was listed first in as many pairs as it was listed second; (*c*) all 190 slips of paper were put into a hat and shuffled, and as they were taken out one by one the order of pairs was made up.

The directions read as follows:

> If we were asked, *"Which is worse, to rob a bank or to steal an apple?"* most of us would say, *"To rob a bank."* If we were asked whether robbing a bank was worse than shooting a man, some of us would say, *"Yes,"* and some would say, *"No."*
>
> In the following examples you are to decide which one of the two deeds printed together is worse and make a cross (X) in front of it. If you aren't sure which is worse, guess. Make one cross for every pair. When you finish the first page go to the second.

2. *The Picture Test*

In a previous paper (9) the writer has described the picture test by which subjects in the first and third grades made their rankings of the offenses. Briefly, it consisted of fifteen 9 by 12 pen and ink outline drawings, reproduced by mimeograph, each illustrating one of the offenses. Five offenses were omitted with these two grades (Offenses 3, 6, 13, 14, and 19) because they were difficult to picture. All the first-grade subjects and 24 of 43 third-grade subjects were tested individually by the writer using these 15 pictured offenses. Each individual test required two sessions of 20 to 30 minutes each. The remaining 19 third-grade subjects were tested in groups of 10 or fewer with the same pictured offenses, but each recorded his own judgments.

Data presented in the earlier paper demonstrated that the statement test and the picture test were comparable. Offense scale-values based on verbal presentation were similar to those based on pictorial presentation; rank order correlation between the two was +.957. The pictures were not complete equivalents for the verbalized offenses, but they were close enough to be worth using with the younger subjects.

3. *Scoring the Paired Comparisons Blanks*

Stimuli arranged in paired-comparisons form and marked by a number of judges can be scored for two purposes. One purpose is to find out how each member of the group ranks the stimuli in question. This involves a simple tabulation of that person's judgments. The other purpose is to find out how the judges as a whole rank the stimuli. This procedure is useful when it is desired to compare the work of these judges with that of another group, or to compare these judges with themselves at a later date.

The object of the present study is to compare boys in one grade with those in another. For this purpose of group comparison Thurstone (27) has provided a useful method. The blanks from each grade must be scored separately. For each grade one must determine the proportion of the group who marked the first offense as worse than the second in every pair of offenses. The resulting table of proportions is then converted into a table of sigma values, and from this table individual scale-values are obtained for each offense. The scale-values use as a unit of measurement the standard error of observation; they reveal not only the rank of the offenses for a particular group, but also how much seriousness separates any two offenses on the scale.

4. *The Interviews*

As a supplement to the offense-rankings, 85 boys from the fifth through the twelfth grades were interviewed. There were from 8 to 12 boys from each of these grades, chosen at random from those who had previously taken the statement test. Below the fifth grade there was increasingly little to be learned by asking subjects why they considered one offense as more serious than another; accordingly, no formal interviews were held with first and third grade boys.

The interviews were standardized and brief. All were conducted by the writer, and no third person was present at any of them. Each boy interviewed was asked to choose the more serious offense in each of 10 pairs of offenses, and was then asked why he judged as he did. The offense pairs used in the interviews were the same for each subject, and are reproduced below:

1. To swipe $1.00 from your brother's bank at home.
5. To lift $1.00 from your father's pants' pocket when taking the pants to the tailor.

10. To keep a ball and glove you find in the school yard.
11. To swipe flowers from a park.

18. To snitch fruit from a peddler's stand.
13. To steal candy and cigarettes from a boxcar.

21. To swipe groceries from your store.
22. To swipe a tie from Marshall Fields.

15. To sneak by an "L" cashier without paying.
20. To sneak a rubber ball from a dime store counter.

6. To keep $1.00 you find on the street, without trying to find the owner.
9. To keep $1.00 you see a man drop from his pocket.

8. To keep a candy package you find after it has fallen from a truck.
2. To help yourself to chocolates from a box in sister's room.

10. To keep a ball and glove you find in the school yard.
16. To take a wagon from a boy's back yard.

5. To lift $1.00 from your father's pants' pocket when taking the pants to the tailor.
17. To swipe a dollar from your boss' desk.

8. To keep a candy package you find after it has fallen from a truck.
14. To ride on the street car for half fare when you should pay full fare.

Each one of these pairs was represented in the test booklet except 21-22; these numbers were given to offenses especially constructed to reveal comparative attitudes toward property owned by a small store and property owned by a large store. The other pairs were chosen because they seemed significant in getting at the relation between the property categories mentioned above.

Effort was made in the interview to get the boy to make as clear a statement as possible of the reasons why one item seemed worse than the other. If the reason he gave bore no logical relation to the way he judged the pair he was questioned until he reached a consistent stand. This was not necessary in most cases. The answers were immediately written down by the interviewer on a prepared sheet. Every precaution was taken against suggesting reasons to the boys interviewed; in the few cases where a boy seemed unable to think of any reason the interviewer let the matter rest there.

D. Results

Results of this study consist essentially of the offense scale-values for each grade and age and the supplementary interviews. Each of these sets of data will be presented separately.

1. *Ranking of Offenses by Grade and Age Groups*

Table 2 contains the scale-values and rank orders for each of the grade groups from 1 through 12. In addition the range of scale-values for each group and the number of boys in each group are listed at the foot of the table. Table 3 contains the same data for groups divided according to chronological age,[2] and includes the adult group as well. The age groupings were made without using boys from Grades 1 and 3 because it was undesirable to combine results from the picture test and from the statement test in one set of scale-values. Age groups 9 and 19 are not included in Table 3 because each contained only six boys. A few members of the adult group overlapped the highest age group, but they were not included there.

Before these data are analyzed and presented graphically certain general considerations should be made clear. In the first place, the two lower grades made judgments on pictured offenses which were at the same time described orally.[3] Data presented in an earlier paper (9) showed that the two forms gave fairly similar results when applied to the same group. Nevertheless it is probable that the shift from individual testing to a group situation, from pictured to verbal stimuli, from the use of two test sessions to the use of one, and from 130 paired comparisons to 190, made results from Grades 1 and 3 less comparable with results from the upper grades than is desirable. An examination of the graphs which follow will show some of the discrepancies that resulted.

At the other end of the age and grade range is the adult group. The group is unfortunately small (32 men). It was chosen in preference to a college student group because of residence in roughly

---

[2]Each age group includes all boys who had passed a particular birthday but had not yet reached the next; e.g., all boys in the 10-year-old group had passed the 10th birthday but had not reached the 11th.

[3]The verbal description of each offense was typed below its picture. The first two or three times each offense appeared in the paired comparisons series the description was read to the subject while he looked at the picture. As soon as he was able to identify a picture without prompting the reading aloud was discontinued.

TABLE 2
Offense Scores and Ranks for Grades 1-12

Grades	1		3		5		6		7		8		9		10		11		12	
Offense Nos.	*S*	*R*	*S*	*R*	*S*	*R*	*S*	*R*	*S*	*R*	*S*	*R*	*S*	*R*	*S*	*R*	*S*	*R*	*S*	*R*
1.	.27	4	.80	5	.41	4	.41	5	.44	4	.46	5	.34	6	.50	4	.48	5	.45	5
2.	—.64	14	—1.17	12	— .71	18	— .85	18	— .64	17	— .81	17	— .92	19	—1.04	19	—1.02	19	— .90	19
3.					—1.02	19	— .87	19	— .94	20	—1.08	20	— .94	20	—1.20	20	—1.25	20	—1.16	20
4.	.10	7.5	.95	3	1.20	1	1.18	1	1.29	1	1.40	1	1.21	1	1.43	1	1.66	1	1.36	1
5.	.57	2	1.20	2	.65	3	.82	2	.75	2	.82	2	.71	3	1.19	2	.81	3	.95	3
6.					.05	11	— .04	11	— .17	12	— .35	15	— .23	14	— .49	15	— .61	16	— .65	15
7.	—.69	15	—1.72	15	—1.03	20	—1.00	20	— .85	19	— .84	18	— .73	17	— .81	17	— .59	15	— .76	17.5
8.	—.35	12	—1.21	13	— .23	15	— .24	15	— .34	14	— .45	16	— .62	16	— .58	16	— .74	17	— .66	16
9.	.35	3	.90	4	.32	6	.40	6	.33	6	.25	8	.38	5	.40	6	.32	6.5	.38	7
10.	—.12	11	— .07	9	— .35	16	— .17	14	— .35	15	— .30	13	— .43	15	— .46	14	— .44	14	— .34	14
11.	—.52	13	—1.47	14	— .20	14	— .14	13	.02	11	— .12	12	.20	10	.15	9	.29	9	.17	9
12.	.23	5	.49	7	.26	7	.21	8	.06	10	.00	10	.30	7.5	.12	10	.18	18	— .05	12
13.					.34	5	.44	4	.38	5	.66	4	.51	4	.35	7	.30	8	.34	8
14.					— .60	17	— .66	17	— .76	18	— .90	19	— .81	18	— .97	18	— .90	18	— .76	17.5
15.	.10	7.5	.43	8	.08	10	— .12	12	— .23	13	— .33	14	— .07	12	— .29	13	— .41	13	— .26	13
16.	—.07	9	.52	6	.17	8	.06	9.5	.29	7	.29	7	.30	7.5	.22	8	.51	4	.40	6
17.	.73	1	1.23	1	.70	2	.60	3	.65	3	.80	3	.77	2	.73	3	1.04	2	.96	2
18.	.16	6	— .26	10	— .06	12	.06	9.5	.13	9	.16	9	.18	11	.48	5	.32	6.5	.51	4
19.					.10	9	.22	7	.16	8	.33	6	.24	9	.06	11	.20	10	.07	10
20.	—.09	10	— .64	11	— .18	13	— .38	16	— .38	16	— .08	11	— .08	13	.05	12	.00	12	.00	11
Range	1.42 σ		2.95 σ		2.23 σ		2.18 σ		2.23 σ		2.48 σ		2.15 σ		2.63 σ		2.91 σ		2.52 σ	
N	56		43		94		94		82		74		88		97		104		104	

Note: Grades are listed along the top, offense numbers down the left side. "*S*" means scale-value, in sigma units; range is from high positive (most serious) to high negative (least serious). Grades two and four were omitted altogether. Offenses 3, 6, 13, 14, and 19 were not pictured, and hence were not judged by grades one and three. "*R*" means rank-order; in grades one and three ranks run from 1 to 15, in all other grades from 1 to 20.

TABLE 3
Offense Scores and Ranks for Age Groups 10 to 18

Ages / Offense Nos.	10 S	10 R	11 S	11 R	12 S	12 R	13 S	13 R	14 S	14 R	15 S	15 R	16 S	16 R	17 S	17 R	18 S	18 R	Adult S	Adult R
1.	.59	4	.28	6	.47	4	.58	4	.41	5.5	.32	6	.51	4	.46	4	.39	8	1.01	3
2.	−.78	19	−.75	18	−.69	18	−.86	17	−.61	17	−.80	18	−1.07	20	−.92	18	−1.12	19	−1.75	19
3.	−.75	18	−.84	19	−.98	20	−1.28	20	−1.01	20	−1.06	20	−.95	18	−.95	19.5	−1.96	20	−1.84	20
4.	1.12	1	1.19	1	1.04	1	1.50	1	1.24	1	1.18	1	1.39	1	1.17	1	1.57	1	1.72	1
5.	.77	2	.66	3	.85	2	.86	3	.68	2	.82	2	.80	3	.89	3	.99	3	[illegible]	4
6.	.08	10	−.04	12.5	.01	11	−.19	13	−.27	15	−.31	14	−.49	15	−.63	16	−1.01	17.5	−1.12	18
7.	−1.05	20	−.93	20	−.88	19	−1.09	19	−.72	18	−.69	17	−.73	17	−.70	17	−.91	16	−.91	15.5
8.	−.44	16	−.27	14	−.37	15	−.43	14	−.31	16	−.33	15	−.69	16	−.55	15	−.78	15	−.91	15.5
9.	.43	6	.46	4	.26	6	.41	6	.41	5.5	.13	10	.29	8	.26	9	.35	9	[illegible]	6
10.	−.18	13	−.34	15	−.43	16	−.46	16	−.26	14	−.41	16	−.39	14	−.40	14	−.43	14	−.15	12
11.	−.34	14	−.02	11	−.14	13.5	−.10	12	.07	10.5	.15	9	.27	9.5	.12	11	−.06	13	−.35	[illegible]
12.	.15	8	.19	7	.16	8	.20	9	.07	10.5	.11	11	.12	12	.21	8	.41	6.5	[illegible]	[illegible]
13.	.36	7	.39	5	.38	5	.48	5	.43	4	.51	4	.34	6	.35	6	.44	[illegible]	[illegible]	[illegible]
14.	−.73	17	−.68	17	−.68	17	−1.03	18	−.55	19	−.94	19	−.97	19	−.98	19.5	−1.01	17.5	[illegible]	[illegible]
15.	−.02	11	−.04	12.5	−.06	12	−.43	15	−.18	13	−.25	13	−.25	13	−.22	11	[illegible]	12	[illegible]	[illegible]
16.	.45	5	.12	9	.18	7	.24	8	.32	7	.29	7	.39	5	[illegible]	7	.61	4	[illegible]	5
17.	.69	3	.73	2	.78	3	.88	2	.54	3	.61	3	.91	2	[illegible]	2	1.16	2	[illegible]	2
18.	−.14	12	.09	10	.04	10	.01	10	.13	8	.36	5	.34	7	[illegible]	[illegible]	.41	6.5	[illegible]	[illegible]
19.	.10	9	.15	8	.14	9	.36	7	.16	9	.19	8	.27	9.5	[illegible]	[illegible]	.23	10	[illegible]	9
20.	−.41	15	−.36	16	−.14	13.5	−.03	11	−.09	12	[illegible]	12	.14	11	[illegible]	12	[illegible]	11	[illegible]	[illegible]
Range	2.17 σ		2.12 σ		2.02 σ		2.78 σ		2.25 σ		2.24 σ		2.46 σ		2.12 σ		[illegible]		[illegible]	
N	27		55		90		89		111		116		124		[illegible]		31		[illegible]	

Note: Ages are listed along the top, offense numbers down the left side. "S" means scale-value, in sigma units; range is from high positive (most serious) to high negative (least serious). "R" means rank order, 1 is most serious, 20 is least serious. Boys from Grades 2 and 3 are not included in this table.

the same area of the city as the younger groups, and because of roughly similar economic status. Two conditions reduce the value of this group for comparison purposes. One is that not all members of the group filled out the questionnaire, and a selective factor is thereby introduced. The other is that the common denominator of the group was unemployment. That unemployment might modify the evaluation of offenses against property is not a fantastic hypothesis, and this may have affected the ranking from that group.

One other aspect of these tabulated data needs to be mentioned. There is no ready way to tell whether the grade grouping or the age grouping should provide the basis for a developmental analysis. If a boy's evaluation of these offenses changes as he grows up, do these changes occur principally as a result of living longer in a certain culture, or are they in some way associated with progress in school? For many boys, of course, gain in chronological age is matched with school progress, but this is not true even for a majority of the boys used in the present experiment. Table 1 shows that the seventh grade included boys from 11 to 17 years of age, and not more than a third were of any single age. Conversely the 15-year age group included boys from Grade 5 through Grade 12, with less than a third from any one grade.

An examination of the scale-values in Tables 2 and 3 shows that there is considerable similarity between the values for grade groups and those for age groups. Since the same judgments are included in both analyses the similarity is not surprising. Because of this similarity, and because the two youngest groups had to be treated as grade groups, the primary analysis in this paper will be made with the data in Table 2, which are based on grade groups.

a. General characteristics of the ranks. It is apparent from an examination of the rank orders of offenses from Grades 1 through 12 that the similarities from grade to grade considerably outweigh the differences. Rank-order correlations between the scale-values of these groups substantiate this. In Table 4 are presented the intercorrelations between ranks for all grade groups and for the adult group. All coefficients are positive; the lowest is .74 and the highest is .99.

As would be expected, correlations between adjacent grades are higher than between grades separated by several years. There is a regular decrease in size of coefficient with increase of separation between grades. In Table 5 the median coefficient for each distance

TABLE 4

Correlations between Rank Orders of Offenses for All Grades and for Adult Group

(All coefficients are *Rho;* for coefficients involving grades 1 or 3 *N* is 15; for all other coefficients *N* is 20; all coefficients are positive.)

	Grades 1	2	3	4	5	6	7	8	9	10	11	12	Adult
1	XX		.89		.88	.87	.82	.80	.82	.79	.74	.75	.80
2		XX											
3	.89		XX		.93	.91	.87	.88	.81	.81	.82	.77	.88
4				XX									
5	.88		.93		XX	.97	.95	.93	.96	.90	.88	.87	.87
6	.87		.91		.97	XX	.98	.94	.96	.91	.88	.88	.84
7	.82		.87		.95	.98	XX	.96	.96	.95	.92	.93	.91
8	.80		.88		.93	.94	.96	XX	.96	.94	.94	.94	.95
9	.82		.81		.96	.96	.96	.96	XX	.95	.94	.92	.92
10	.79		.81		.90	.91	.95	.94	.95	XX	.98	.99	.96
11	.74		.82		.88	.88	.92	.94	.94	.98	XX	.98	.97
12	.75		.77		.87	.88	.93	.94	.92	.99	.98	XX	.96
Adult	.80		.88		.87	.88	.91	.95	.92	.96	.97	.96	XX

TABLE 5

Relationship between Size of Inter-Grade Coefficients and Number of Years between Grades

(Computed from coefficients in Table 4; adult group not counted as a grade.)

Distance between grades	Number of coefficients	Median coefficient
1 year	7	.97
2 years	8	.94
3 years	6	.935
4 years	6	.915
5 years	5	.88
6 years	4	.85
7 years	3	.81
8 years	2	.82
9 years	2	.78
10 years	1	.74
11 years	1	.75

between grades shows this progression clearly. Figure 1 shows graphically the relation between size of intergrade coefficient and number of years between grades.

Two generalizations seem indicated by these results. (*a*) Since all correlations are positive, and all significantly above zero, it follows that whatever capacities are required to judge the seriousness

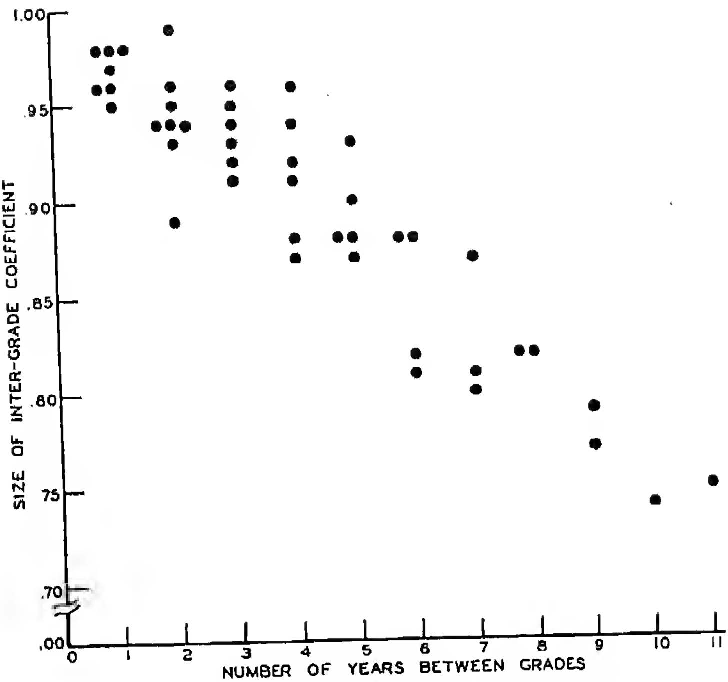

FIGURE 1

RELATION BETWEEN SIZE OF INTER-GRADE COEFFICIENTS AND NUMBER OF YEARS BETWEEN GRADES

of property offenses are present in many boys of six, and that major changes in the evaluation of these offenses occur infrequently thereafter. (*b*) Since similarity in age and grade status is paralleled by similarity in evaluation of offenses, and age and grade differences parallel decreasing similarity in judgments, it follows that changes which do occur from grade to grade in the ranking of offenses are not haphazard and chance, but are regular and perhaps predictable. These slight but regular changes may represent progressive approximations of an adult social norm which is already accepted in ambiguous and incomplete fashion by first-grade boys.

Table 4 requires further comment. There are 19 coefficients which involve either the first or the third grade. Only two of these are above +.90. There are 36 coefficients which do not involve

either of these two grades, and 30 of these are +.90 or above. The median coefficient of these 36 is +.94. There is evidence here that changes in the ranking of offenses from the fifth grade on are very slight, and that whatever concept of property rights is responsible for these judgments is relatively stable in that time.

It should be noted here that high correlations between the offense rankings of similar groups is not unusual. Brogan (3, 4, 5, 6) and Weber (31), in a series of studies using the 16 "worst practices" among university students, found correlations averaging +.95 between rankings of high school and college men and women in various parts of the country. The paired-comparisons method was not used in these studies.

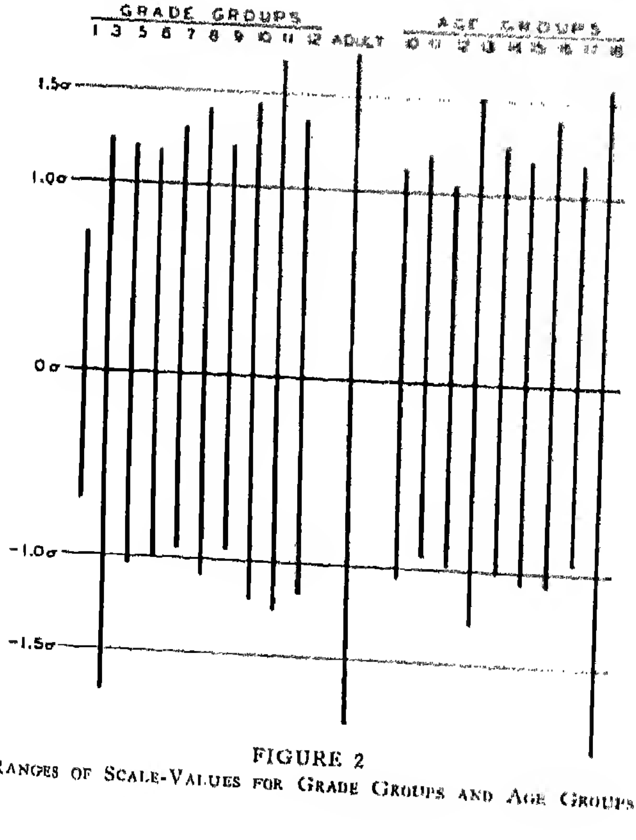

FIGURE 2
RANGES OF SCALE-VALUES FOR GRADE GROUPS AND AGE GROUPS

b. Differences in range of scale-values for the various groups. Tables 2 and 3 list at the bottom of each column of scale-values the range in values for that group. Figure 2 shows these ranges graphically.

The range for Grade 1, 1.42 σ, is smaller than that of any other group. The adult group with a range of 3.56 σ is largest. The median range for all groups is 2.35 σ.

It is probable that ordinary fluctuations in these ranges are not of very great significance, serving only to indicate some of the chance factors operating in the data. Large differences, however, deserve consideration.

The factor which determines the range of values for any group is the degree of unanimity shown by the group in judgments on the paired stimuli. If a group were split 50-50 on every paired comparison, the resulting range would be zero. If a group were unanimous in judging the more serious offense of every pair, the resulting range would approximate 6 sigma with the method used here. Thurstone has said, in a study of nationality preferences, "A wide range of scale values indicates rather strong and rather uniform national preferences" (28). In this study it may be said that, other things being equal, a wide range of scale values indicates rather strong and rather uniform opinions about the seriousness of property offenses.

It is possible to show empirically the relationship between range of scale-values and amount of agreement within the group. The paired-comparisons blank can be scored to obtain the offense ranking of the individual subject as well as the group scale-values. To obtain the former it is necessary to tabulate for each paper the number of times each offense is checked as worse than those with which it is paired. These tabulations can then be arranged in a rank order of seriousness for that subject. If there is much agreement among the members of a group the individual rank orders should correlate highly, on the average, with the rank order of scale values for the group. If there is little agreement the correlations between individual ranks and the group rank should be low. When individual ranks and correlations with the group rank were computed for three groups, the results listed in Table 6 were obtained.

These data show that there actually was much closer agreement in the adult group than in either age 14 or age 11. The sizes of

TABLE 6

RANK-ORDER CORRELATIONS BETWEEN INDIVIDUAL OFFENSE RANKS AND GROUP RANKS

Group	N	Mean correlation (rho)
Adult	32	+.831
Age 14	111	+.705
Age 11	55	+.706

the mean coefficients correspond with the ranges of offense scale-values.

There is a fatigue and boredom error in any paired-comparisons rating which must not be ignored. That error is maximized when the paired items are complex, and minimized when the paired items are simple. The offense descriptions used for most of the subjects in this experiment were considerably more complex than the single-word items usually employed in the paired-comparisons procedure. The stated offenses average 10 words in length, and the subject who took the statement test had to read each statement 19 times. That makes 380 statements, and approximately 3,800 words of reading matter. For schoolboys of fifth-grade status and above that is not too difficult a task, but the large amount of repetition involved could, and probably did, make the task an irksome one. Boredom, fatigue, and irritation would add to the chance errors in judging, and would tend to reduce the amount of agreement among the judges. This would narrow the range of scale-values. It should be said here that observable irritation among the subjects in this experiment while they were marking the paired-comparisons blank was very slight, but the inference from the nature of the blank is nonetheless sound.

In a sense the pictured offenses with their verbal descriptions are also complex stimuli, but in another sense they are much less so than are the statements alone. None of the subjects tested individually with the picture test read the descriptions underneath the pictures. Each of these subjects heard the descriptions read to him by the experimenter during the early part of each of his experimental sessions, but it is safe to say that for him the stimulus was a meaningful picture. The meaning of the picture was clarified by the description, but the description itself could be ignored after the picture became sufficiently familiar. Since pictures are more concrete than words they restrain the wandering attention and they minimize

boredom. The use of pictures would thus prevent error from concealing real agreement, and would tend to widen the range of scale-values. Many of the first- and third-grade boys actually seemed to enjoy the sessions with the pictures.

One additional factor affecting the unanimity of judgments is the degree of homogeneity in the group. If each grade group had been restricted to boys who were of normal age for their grade the range of scale-values for each group would undoubtedly have been wider, because the groups would have been more homogeneous. Heterogeneity, in age or grade, would reduce agreement within the group and hence narrow the range of scale-values.

In the light of these facts the differences in the range of scale-values for the age and grade groups become less puzzling. For the most part the differences are slight, but there are certain outstanding exceptions. The two widest ranges are the adult group and the 18-year-old group, with 3.56 σ and 3.53 σ respectively. These are also the two smallest groups represented in Figure 1, with 32 and 31 subjects respectively. All but four boys from the 18-year-old group came from Grade 12, thus insuring considerable grade homogeneity. The adult group was not a random sampling, for it represented a 43 per cent return on a mailed questionnaire, and this selective process may have contributed to the homogeneity of the group.

Unusually wide ranges appeared also in the third grade, the eleventh grade, and the 13-year-old group, where they were 2.95 σ, 2.91 σ, and 2.78 σ respectively. There is no obvious explanation for the wide range of the last two of these groups, for both were typical in size (104 and 89 subjects respectively) and in heterogeneity. The third grade had best be considered in connection with the first grade, since both used the picture-test procedure.

The range of first-grade scores is by far the narrowest of all the groups studied. This is true in spite of the fact that each subject was tested individually with pictures, that before scale-values were calculated 14 out of 70 subjects were eliminated because of failure to master the test procedure, and that the group was more homogeneous in age than most of the grade groups (35 boys were 6 years of age, 17 were 7, and 4 were 8). All these factors tend to widen the range of scale-values, and yet the range for Grade 1 is extremely narrow. The conclusion must be, therefore, that first-grade boys disagreed markedly in their estimation of the seriousness of the

offenses used. In view of the age and homogeneity of the group these disagreements probably represent weak and uncertain evaluation of the offenses, rather than conflicts between strongly held opinions; if this is true the judgments of the first grade mark an early stage in the developing awareness and acceptance of a hierarchy of values in the property area.

The third grade, in contrast to the first and to the next higher grade studied, the fifth, has a wide range of values. The third grade is the smallest group of the three, is about as homogeneous in age as the first grade, and considerably more so than the fifth grade. Since the third grade was tested under almost the same conditions as the first grade, and was equally homogeneous with respect to age, it is inferrable from a comparison of the ranges of scale values that the third grade had reached a markedly greater agreement on the seriousness of the offenses. If we regard the changes from grade to grade in offense rankings as progressive approximations of an adult social norm, it seems clear that the change from Grade 1 to Grade 3 constituted a considerably larger step in this direction than did the change between any other two grades. Presumably, then, the first two or three years in school constitute a period of particularly rapid adaptation to a compelling social norm.

These statements assume that if the fifth and higher grades had been tested individually with the picture test they would have shown higher internal agreement than they actually did because of the elimination of experimental errors. In this case the difference in range between Grade 3 and the higher grades would disappear. This seems to be a reasonable assumption.

c. Grade differences in judging the seriousness of the offenses. The principle results of this study have already been stated. They are, briefly, that property offenses can be ranked in seriousness by first-grade boys, that at this level differences between the offenses are confused and indistinct, that between the first and the third grade rapid progress is made in evaluating the seriousness of offenses, that above the third grade changes take place much more slowly and more infrequently, that similarities in offense evaluations between groups are directly proportional to similarity in age and grade status, and that changes in offense evaluation from grade to grade are not haphazard but are regular and orderly as if gradually approaching an "ideal" or a "correct" evaluation. There remains the task of

examining the nature of the changes that take place in the ranking of specific offenses.

The 20 offenses used in this study are of little significance in themselves. They were chosen because they represented fairly frequent experiences in the lives of these Chicago boys, but another list might have served the purpose. Certainly different groups of boys living in contrasting situations would require a new list of offenses if a study similar to this were to be made. The offenses used here cannot be thought of as more basic or more fundamental than an alternative list except as they are more familiar to the boys who judged them. Value judgments such as these are relevent to a time, a place, and a person. They have no eternal, universal significance. Thus the degree of seriousness assigned to specific offenses is not of major importance here. There is obviously no final or correct order of seriousness in any absolute sense, even though this study seems to show a progression from the first through the twelfth grade in the direction of *an* order. Nevertheless there may be some interest in examining what happens to each offense as we pass from grade to grade.

The relative seriousness of each offense at each of the grade levels and for the adult group is shown in Figures 3, 4, 5, and 6 below.

In Figure 3 are plotted the scale-values for four offenses. Grade groups are listed along the top, and the sigma scale of seriousness along the left side. High positive values indicate the most serious offenses, high negative values the least serious. Each scale-value is indicated by a small circle. From the fifth through the twelfth grade these are connected by solid lines, for these grades are thoroughly comparable. Lines connecting the two lower grades and the adult group are broken because the comparability of these groups is not thoroughly established. In particular the extended range of the values in Grade 3 diminishes the graphic similarity they might otherwise have had to Grades 1 and 5.[4]

An examination of the plotted scale-values showed that whether an

[4]The range of Grade 3 could have been adjusted by a simple reduction process to one that approximated the ranges of the other groups. Thurstone (29) recommends that ranges be so adjusted when comparing two sets of scale-values from the same judges. In the present instance, however, different judges are involved, and the extended range of grade three seems to be a direct result of the method used. It has therefore seemed preferable to plot the values without distortion, even though the resulting graph is less clear.

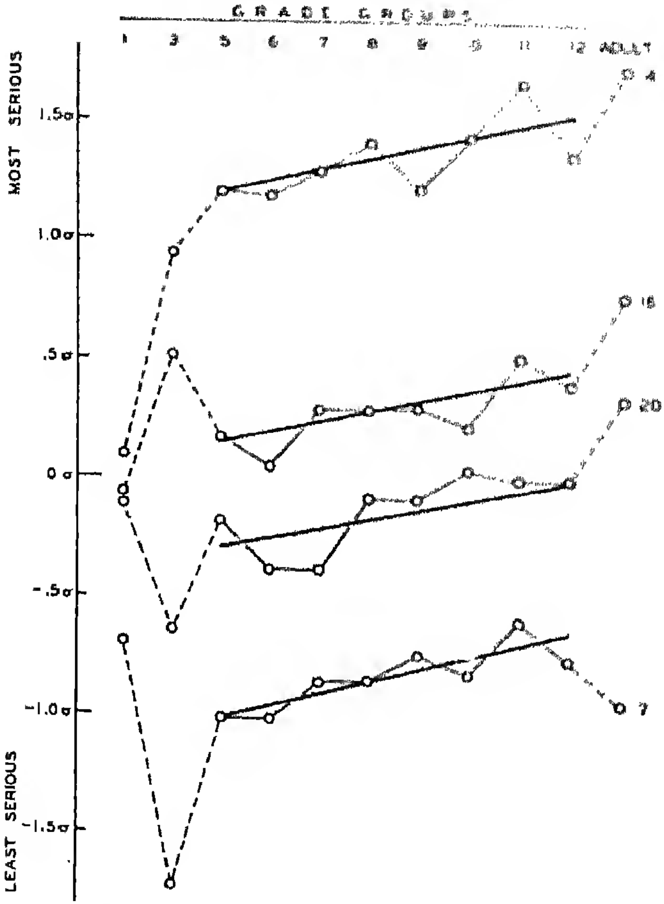

FIGURE 3
SERIOUSNESS OF OFFENSES 4, 7, 16, AND 20 FOR ALL GRADES AND THE ADULT GROUP

offense gained or lost in seriousness through the upper eight grades, the plotted values could in most cases be best represented by a straight line. Accordingly, a straight line has been fitted to each offense, only the values for the upper eight grades being used. The lines were fitted by eye and are approximate only. A glance at the

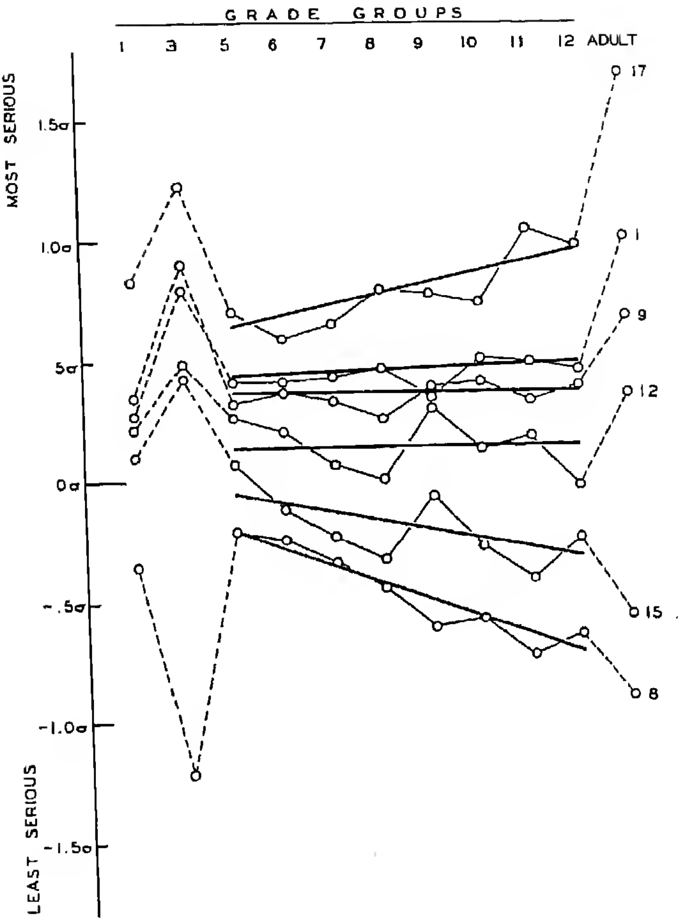

FIGURE 4

SERIOUSNESS OF OFFENSES 1, 8, 9, 12, 15, AND 17 FOR ALL GRADES AND THE ADULT GROUP

graphs will show that some offenses yield to this treatment better than others do.

The trend of each of the four offenses in Figure 3 is toward greater seriousness. Offense 4, *"to swipe your mother's wrist watch and pawn it,"* was the most serious offense for all grades above the

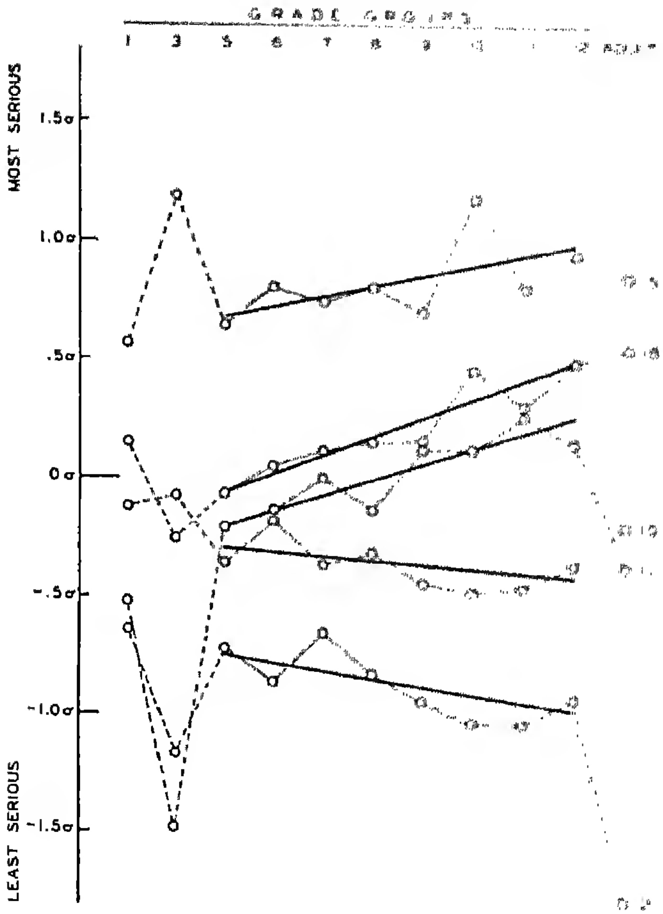

FIGURE 5
SERIOUSNESS OF OFFENSES 2, 5, 10, 11, AND 18 FOR ALL GRADES AND THE ADULT GROUP

third, including the adult group. Offense 16, "*to take a wagon from a boy's backyard,*" is ranked variously from fourth to ninth in seriousness. Offense 20, "*to sneak a rubber ball from a dime store counter*" is ranked tenth to sixteenth in seriousness. Offense 7, "*to take a wheel from a wagon you find in the alley*" is ranked from

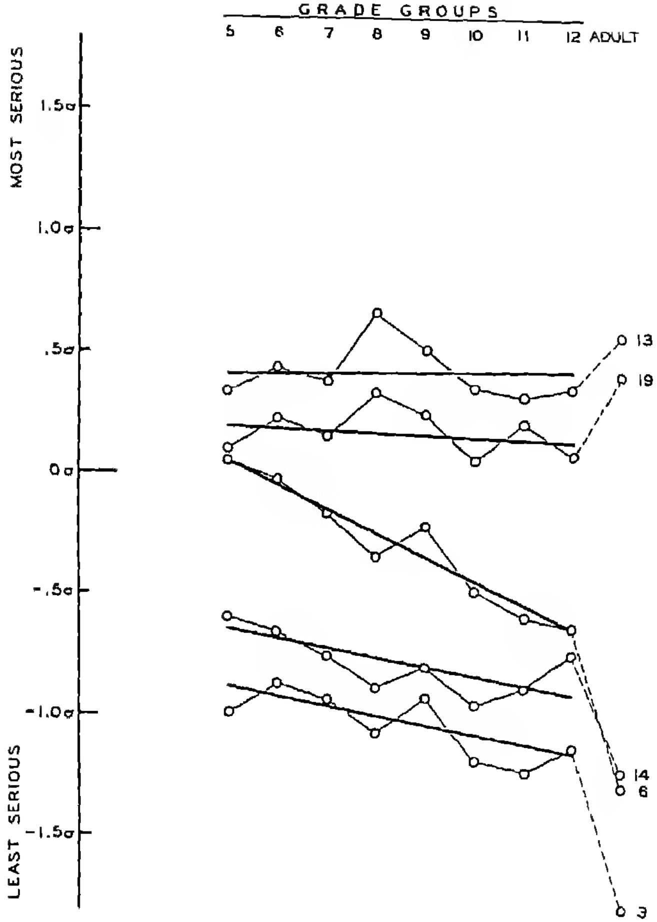

FIGURE 6

SERIOUSNESS OF OFFENSES 3, 6, 13, 14, AND 19 FOR ALL GRADES AND THE ADULT GROUP

fifteen to twentieth. The first four grades regard this as the least serious of all. With older groups it gradually becomes more serious, until for the eleventh grade there are five less serious offenses.

The six offenses plotted in Figure 4 are concentrated fairly well in the middle range. Three make no significant change in serious-

ness, two get less serious in the higher grades, and one gets more serious. The latter is Offense 17, *"to swipe a dollar from your boss's desk."* The two which get less serious are Offense 15, *"to sneak by an 'L' cashier without paying,"* and Offense 8, *"to keep a candy package you find after it has fallen from a truck."* It is of interest to note that Offense 17 is judged by the adult group as almost as serious as Offense 4; all other groups above the third grade judged Offense 4 as considerably more serious.

Five more offenses are represented in Figure 5. One, Offense 10, does not change much in seriousness. It reads, *"to keep a ball and glove you find in the school yard."* Three become more serious. They are Offense 5, *"to lift $1.00 from your father's pants' pocket when taking the pants to the tailor"*; Offense 18, *"to snitch fruit from a peddler's stand"*; and Offense 11, *"to swipe flowers from a park."* Offense 5 is second or third in seriousness for all grades, but gains in seriousness in that scale separations are wider in the higher grades. Offense 18 gains in seriousness not only through broader scale separations, but also in comparison with other offenses. It is ranked twelfth in seriousness by Grade 5, and fourth by Grade 12. Offense 11 increases in seriousness rank from fourteenth to ninth, although the adult group apparently considers swiping flowers from a park a much less serious offense than do the higher grades.

Offense 2 becomes slightly less serious in the higher grades, although not so in comparison with other offenses. It reads, *"to help yourself to chocolates from a box in your sister's room,"* and is ranked from seventeenth to nineteenth by all groups. Through all these graphs there runs a tendency for scale separations between offenses to be greater in the upper grades than in the lower; that is, distinctions in seriousness between the offenses are more clearly seen by the older groups. An offense may therefore gain or lose in seriousness score without changing its rank position in the list of offenses.

The offenses plotted in Figure 6 are those which were not included in the picture test. Accordingly there are no scores plotted for Grades 1 and 3. Offense 6, *"to keep $1.00 you find on the street, without trying to find the owner,"* is the only one in Figure 3 which undergoes marked change. It becomes much less serious in the higher grades. It was ranked eleventh in seriousness by the fifth grade, and eighteenth by the adult group.

In summarizing the plotted data it will be useful to separate the offenses into three groups on the basis of the amount and direction

of change through the upper eight grades. The classification below is based on the trend lines drawn to fit the plotted data. The first group includes those offenses which become more serious in the upper grades, listed in order of the magnitude of the change. The offense which changed most is listed first. The second group includes those offenses that became less serious, again listed in order of the magnitude of the change. The third group includes the offenses which did not show a significant change.

Offenses which gained in seriousness

18. To snitch fruit from a peddler's stand.
11. To swipe flowers from a park.
7. To take a wheel from a wagon you find in the alley.
4. To swipe your mother's wrist watch and pawn it.
5. To lift $1.00 from your father's pants' pocket when taking the pants to the tailor.
17. To swipe a dollar from your boss's desk.
16. To take a wagon from a boy's back yard.
20. To sneak a rubber ball from a dime store counter.

Offenses which lost in seriousness

6. To keep $1.00 you find on the street without trying to find the owner.
8. To keep a candy package you find after it has fallen from a truck.
2. To help yourself to chocolates from a box in your sister's room.
3. To borrow your brother's baseball without asking.
14. To ride on the street car for half fare when you should pay full fare.
15. To sneak by an "L" cashier without paying.

Offenses which did not change in seriousness

1. To swipe $1.00 from your brother's bank at home.
9. To keep $1.00 you see a man drop from his pocket.
10. To keep a ball and glove you find in the school yard.
12. To snitch three tickets from a movie cashier.
13. To steal candy and cigarettes from a boxcar.
19. To swipe and sell lead pipes from an old warehouse.

Some comments can be made about these classifications. The offenses which gained in seriousness all involve actual stealing in one form or another. The three offenses that were consistently

the most serious of all (4, 5, 17) are in this group, and only two out of the eight offenses in this category fall in the less serious half of the scale. In other words, the serious offenses (with the exception of some which did not change) became more so.

In the second category it is to be noted that hoodwinking the utilities, keeping found property, and using the belongings of sibs in the home are regarded less seriously by the older groups. All of these offenses fall in the less serious half of the scale; with minor exceptions all scale-values in this group are negative. In other words, the less serious offenses became even less so. The one exception to this statement is Offense 7, *"to take a wheel from a wagon you find in the alley,"* which is the least serious of all for the lower grades, but which becomes more serious in the older groups. Hypotheses may be advanced to explain this, but there is no one obvious answer.

The offenses which did not change significantly in seriousness are located for the most part in the middle range of the scale. None is among either the three most serious or the three least serious offenses. The absence of change in some of these offenses, in the light of changes in apparently similar ones, is hard to understand. Why should stealing from father or mother gain in seriousness, while stealing from brother does not change? Why should stealing from the dime store or a peddler gain in seriousness when stealing from a movie, from a boxcar, and from a warehouse do not change?

It seems that apparent similarity may conceal significant differences in these offenses. A description of these hidden differences would require further research.

d. Results of the interviews. Interviews with 85 boys chosen about equally from the upper eight grades threw some light on the criteria for judgments of seriousness. In the interview each boy judged 10 pairs of offenses and gave his reason for each decision. These reasons make an interesting study.

It appears from the interview data that an offense will be judged as more serious in proportion as it (*a*) seems more likely to bring punishment, (*b*) involves property of greater value, (*c*) is owned by one toward whom obligation is felt (e.g., mother), (*d*) injures an actual person, (*e*) must be called "stealing" rather than a more euphemistic term such as "borrowing," or (*f*) damages the moral character of the offender. Contrariwise the absence of these features will render an offense less serious. An offense will also be judged

less serious if (*a*) the owner of the property is unknown or hard to locate, (*b*) the property can be conceived of as sharable (e.g., sister's chocolates), or (*c*) the property is a service, such as an "L" ride, and not a tangible object.

All these reasons and a number of others were offered by the boys interviewed. Some appeared predominantly in one part of the grade range and some in another. Perhaps the most striking difference between grade groups was in the variety of reasons offered. The higher grades were notably more ingenious in furnishing a varied list. Another significant grade difference was in the frequency with which fear of punishment was offered as a reason. In the fifth grade 53 per cent of the reasons given were based on a fear of punishment; in the twelfth grade only 3 per cent of the reasons were in this category. Intervening grades varied roughly in proportion. These data do not support the general statement of McGrath (20), based on asking boys and girls of 6 to 18 why stealing is wrong, that "But a very small percentage of the cases considered give as the reason why this act is wrong that 'they might get caught.' The attitude of the average child is, therefore, the correct one and differs materially from the unwholesome one so often found in the delinquent child." Perhaps McGrath's subjects, most of whom were from parochial schools, learned sooner than the boys in this study that fear of punishment is not a respectable reason for condemning stealing.

A third difference between grades was the increasing frequency with which boys in the upper grades judged the seriousness of an offense in terms of who was hurt and how much. This concern with the effects of an act in contrast to judgments based on some pseudo-absolute moral standard, seems to be characteristic of the increasing social maturity that is usually a part of the process of growing up. Many studies of "moral judgment" have been made in which the subjects were asked to assign praise or blame to the hero of a fictional incident. Virtually all of these studies have shown that with increasing age there is increasingly greater reliance placed on the motive for the act as a criterion of its moral worth.[5] Since mention of motives was intentionally excluded from the offenses

[5]Cf. Schallenberger (24), Schaefer (23), Kolb (17), Levy-Suhl (18), Piaget (22), Barnes (1), Jacobsohn-Lask (15), Tudor-Hart (30), Slaght (26), Clem and Smith (8), Hecker (12).

used in this study, the same tendency seems to have been expressed in another way—by a consideration of effects.

It must be recognized that the task of furnishing reasons for judgments of offense seriousness is one which places a premium on social intelligence. The interview makes a situational demand that the reason offered be as "good" as possible. Not all boys respond equally to this demand, but most of them undoubtedly are influenced by it. Fear of being caught is a real deterrent to many of the boys used in this study. But their decisions about offenses were verbal answers to verbal situations, and on the verbal level fear of being caught is not acceptable to moralizing adults as a basis of moral judgments. The choice of reasons throughout the grade range reflects a growing awareness of what is acceptable to adults in authority.

E. Discussion

Instead of describing as "moral" or "intellectual" the growth depicted by the results of this study, it seems better to speak of "social growth." There is evidence here that a practical discrimination which in effect recognizes the existence of property rights develops early, and is in workable shape at the 6-year-old level. This development would not take place in the absence of "mental" growth, nor in the presence of a markedly retarded mental growth. Furthermore it would not take place in the same way in a society which valued property differently.

The evidence furnished by the offense ranking procedure shows how a group of boys in a large city are conditioned to behave in accordance with accepted social norms. The behavior sampled here is verbal behavior, but that is at least as closely regulated by social norms as is any other kind of behavior. Property norms are especially pressing, and the child must become aware of them early if he is to reduce conflict over property to a minimum. It must be remembered too that the offense ranks of the youngest group of subjects tested were in general outline very similar to the ranks of the older groups tested. In other words the experiences of the 6-year-old child of the economic and social class used in this study have in most cases been sufficient to enable him to rank property offenses much as do boys 10 to 15 years older. It is evident that certain fundamental distinctions are established early in the life of the child, and that these distinctions remain effective through boyhood and adolescence.

How early are these distinctions actually arrived at, and how does this phase of development compare with growth of other social attitudes? There is not much evidence on the first question, although general studies of infancy and childhood have furnished some clues. It seems probable that before three years of age there is no feeling of property rights at all in the child. For the first year or two of life behavior is largely impulsive. The child's experience with material objects during this time is necessarily limited. He seems to be quite indifferent to property distinctions, making use of whatever will satisfy his needs of the moment. But as soon as he becomes old enough to play with other children—between two and three years of age—the desire for possession appears with considerable intensity. Isaacs (14) describes this type of response as unlearned because of the spontaneity of the reactions and their "toughness under training." The desire for possession is fundamentally a social one, having little relation to the intrinsic value of the property. The child wants the ball not because of a specific desire for a ball, but because some other child has it.

Beaglehole (2) states that from about the age of three years the child begins to get some hint of the distinction between his own possessions and those of other individuals. Healy (11) says that "at 3 or 4 years one can expect little perception and controlled conduct," but that before 6 years there is usually to be found development of a sense of property rights "adequate for simple needs." How this feeling arises is hard to say, but it doubtless depends a good deal on the inhibitions set up by means of adult restrictions. France and Kline (16) report that the adults who answered their questionnaire about childhood experiences with property developed a sense of property rights "by making objective those feelings of care of property, love of possession, pride in ownership" which characterized their own relation to property. This process of projecting one's feelings onto other individuals, and realizing that they too may feel as we do, seems a little sophisticated. The adults in this study may have failed to recall their childhood experience accurately, although some process of this sort probably takes place sooner or later with most people. What happens to property attitudes before the age of 6 needs further study before the whole story of growth in this respect can be told.

The early presence of evaluative attitudes toward property offenses is paralleled by the development of some other social attitudes.

Race and nationality prejudices have been found by several investigators to begin early. Minard (21) found the race attitudes of Iowa children to be fairly stable by the time the seventh grade was reached. He used no subjects younger than this and could not tell how much earlier these prejudices began. After the seventh grade changes were infrequent. Green (10), using children from 7 to 16 in Wales, found prejudice against Chinese, Negroes, French, Germans, Italians, Spaniards, Americans, and Russians already fixed at the age of 7, and no age differences except in the ingenuity with which opinions were defended. Horowitz (13) observed prejudice against Negroes in northern and southern children at the lowest level tested, barely over 5 years of age. Incidental observations on 3- and 4-year-old boys also gave evidence of the existence of prejudice. In a different area Lockhart (19) found that the attitudes toward law of children from grades 4 to 12 change but gradually toward an adult norm. McGrath (20) discovered that 6-year-old boys could diagnose a stealing situation as "stealing" with about the same facility as adults. Carmichael (7) found that a large proportion of 6-year-old children recognize "truth" as a desirable aspect of behavior.

All of these data seem to indicate the rather early development and general stability of certain basic social attitudes of the child. The attitude toward property, in the sense of habit patterns and concepts and sentiments which do not permit the free play of all impulses regarding property, probably develops as soon as any, because it is necessary for the child's successful adjustment to society.

It appears from the evidence gathered in this study that attitudes toward property are not simple, and that any concept of property rights is to a considerable extent fluid and conditional. A violation of property rights is not, of itself, necessarily serious. Among the relevant factors which determine its seriousness are the relationship of the owner to the offender, the danger of punishment for the offender, the likelihood of real injury to the owner, the kind of property involved, and the value of the property. There may be others. Insight into these factors undoubtedly is greater in the higher grades. This may explain some of the grade changes found in ranking the offenses.

F. Conclusions

1. Attitudes toward property, as indicated by a paired-compari-

sons ranking of offenses, begin to achieve stability at about the six-year-old level for boys of the social and economic class used in this study. Judgments given by boys of this age indicate that, although little unanimity is present, the group scores on the offenses are similar in rank order to the scores of older groups. The lack of agreement between members of the first grade group indicates that the concepts of "seriousness" are very much in the formative stage.

2. Between the first and the third grades rapid progress is made in evaluating the seriousness of offenses and in achieving agreement within the group on offense seriousness. Above the third grade change in this direction continues, but more slowly.

3. Similarity in offense evaluation between grade groups is inversely proportional to distance between the grades. Offense ranks of adjacent grade groups correlate on the average +.97. More widely separated grades give lower coefficients.

4. Changes which occur from grade to grade in the ranking of offenses are regular and orderly, and suggest that with increase in age and grade there is a progressively closer approximation of an adult social norm.

5. In general there were wider scale separations between the offenses in the upper grades, indicating a clearer perception of distinctions in seriousness. For this reason serious offenses tended to become more serious for older groups, and offenses at the other end of the scale tended to become even less serious. Certain exceptions were noted.

6. The a priori classifications of offenses into the four categories of home property, lost property, property with many owners, and property with one owner did not prove to be a serviceable one. Neither in place on the seriousness continuum nor in tendency to change in seriousness with age did all the offenses in one category behave alike. No classification of offenses in terms of one variable alone would be very meaningful to the subjects used in this study.

7. In general younger boys gave the fear of punishment more frequently as a reason for judgments of the seriousness of offenses than did older boys. In the upper grades the reason given most frequently was an unwillingness to injure others. Older boys differed also in that they presented a much wider variety of reasons than did younger boys.

8. Judgments of the seriousness of property offenses seem to be based principally upon the relationship of the owner to the offender,

the likelihood of real injury to the owner, the possibility of punishment of the offender, the kind of property involved, and the value of the property. Another influential criterion, particularly with older subjects, would be the offender's motive, if it were included in the description of the offense.

References

1. Barnes, E. Growth of social judgment. *Barnes' Studies in Education*, 1902, **2**, 203-217.
2. Beaglehole, E. Property: A Study in Social Psychology. New York: Macmillan, 1932.
3. Brogan, A. P. A study in statistical ethics. *Internat. J. Ethics*, 1923, **33**, 119-134.
4. ———. Group estimates of frequency of misconduct. *Internat. J. Ethics*, 1924, **34**, 254-271.
5. ———. Moral evaluations about men and women. *Internat. J. Ethics*, 1925, **35**, 105-124.
6. Brogan, A. P., & Slavens, G. S. Moral judgments of high school students. *Internat. J. Ethics*, 1927, **38**, 57-69.
7. Carmichael, A. M. Moral judgments of first-grade children and a method of securing them. *Elem. School J.*, 1929, **29**, 681-690.
8. Clem, O. M., & Smith, M. Grade differences in attitudinal reactions of six-year secondary school pupils. *J. Educ. Psychol.*, 1934, **25**, 297-309.
9. Eberhart, J. C. The use of pictures in the estimation of the seriousness of property offenses. *J. Genet. Psychol.*, 1940, **56**, 411-437.
10. Green, G. H. Have children a national bias? *Discovery*, 1932, **13**, 44-46.
11. Healy, W. Honesty. Indianapolis: Bobbs-Merrill, 1915.
12. Hecker, W. Über die sittliche Entwicklung von Schulkindern und Frühjugendlichen. *Neue psychol. Stud.*, 1937, **7**, 5-148.
13. Horowitz, E. L. The development of attitude toward the Negro. *Arch. of Psychol.*, 1936, No. 194.
14. Isaacs, Susan. Social Development in Young Children. London: Routledge, 1933.
15. Jacobsohn-Lask, L. Über die Fernald'sche Methode zur Prüfung des sittlichen Fühlens. *Beiheft. Zsch. f. angew. Psychol.*, 1920, **26**, 1-84.
16. Kline, L., & France, C. J. The psychology of ownership. *Ped. Sem.*, 1899, **6**, 421-470.
17. Kolb, E. Die sittliche Entwicklung des Heranwachsenden im Lichte der exacten Forschung. *Päd. Monog.*, 1922, **20**.
18. Levy-Suhl, M. Die Prüfung der sittlichen Reife Jungendliche angeklagter, etc. *Zsch. f. Psychiat.*, 1912, **4**.
19. Lockhart, E. G. The attitudes of children toward law. *Univ. Iowa Stud. Char.*, 1930, **3**, No. 1.

20. McGrath, M. C. (Sister Mary). A study of the moral development of children. *Psychol. Monog.*, 1923, **32**, No. 144.

21. Minard, R. D. Race attitudes of Iowa children. *Univ. Iowa Stud.: Stud. Char.*, 1931, **4**.

22. Piaget, J. The Moral Judgment of the Child. New York: Harcourt Brace, 1932.

23. Schaefer, M. Elemente zur moral-psychologischen Beurteilung Jugendlicher. *Zsch. f. Päd. Psychol.*, 1913, **14**, 47-59, 90-98.

24. Schallenberger, M. E. A study of children's rights, as seen by themselves. *Ped. Sem.*, 1894, **3**, 87-96.

25. Shaw, C. R. Delinquency Areas. Chicago: Univ. Chicago Press, 1929.

26. Slaght, W. E. Untruthfulness in children: its conditioning factors and its setting in child nature. *Univ. Iowa Stud.: Stud. Char.*, 1928, **1**, No. 4.

27. Thurstone, L. L. The method of paired comparisons for social values. *J. Abn. & Soc. Psychol.*, 1927, **21**, 384-400.

28. ———. An experimental study of nationality preferences. *J. Gen. Psychol.*, 1928, **1**, 405-425.

29. ———. Influence of motion pictures on children's attitudes. *J. Soc. Psychol.*, 1931, **2**, 291-305.

30. Tudor-Hart, B. E. Are there cases in which lies are necessary? *Ped. Sem.*, 1926, **33**, 586-641.

31. Weber, C. O. Moral judgments in female delinquents. *J. Appl. Psychol.*, 1926, **10**, 89-91.

University College
Northwestern University
Chicago, Illinois

The Journal of Genetic Psychology, 1942, **60**, 37-60.

THE BODY IMAGE IN ADOLESCENT BOYS*[1]

Bellevue Hospital, New York City

FRANK J. CURRAN, M.D., AND J. FROSCH, M.D.

According to Schilder (13) "the image of the human body means the picture of our own body which we form in our mind, that is to say the way in which the body appears to ourselves." The body image has come in for recent consideration by L. Lhermitte (9). In an attempt to study the body image and its concepts in the adolescent, we examined a group of 100 boys ranging in ages from 12 to 16 who were patients on the Adolescent Ward at the Bellevue Psychiatric Hospital. This ward handles adolescent boys from all parts of New York City—the majority of whom are non-psychotic behavior disorders. A description of the organization of this ward and the types of problems treated there has been published elsewhere (4).

In the course of the study, we observed from time to time cases whose clinical picture was dominated by an excessive interest in the body image and parts of the body image. The diagnosis and classification of these cases presented some difficulty. We were at a loss to evaluate the significance of this excessive body preoccupation as we had very little to guide us as to the extent to which the average adolescent is preoccupied with his body or shows interest in the different parts of his body. The report of David Levy (7) in which he studied the reaction of children to the different parts of their body was worked out mainly on children much younger than the age group met with on our service. We could not help but feel that puberty which brought with it such a tremendous change in the anatomic and physiologic structure of the body with its consequent emotional reactions would also bring with it a somewhat modified attitude towards the body. The beginning differentiating of the secondary sex characteristics, the rapid growth of all parts of the body, the concomitant emotional upheavals are all bound to have their repercussion and expression in the adolescent's attitude

*Received in the Editorial Office on September 2, 1940.

[1]Read before American Psychiatric Association, May, 1940. From the Psychiatric Division of Bellevue Hospital and the Department of Psychiatry, New York University Medical College, New York, N. Y.

toward the body. Wile (16) expresses this attitude when he says—"No period of years except the early years of life is more significant for physical and physiological development than adolescence which marks the transforming potency of gonadal activity. . . ." He goes on to say:

> The rapid growth of bone and muscle, the oscillation between energy production and fatigue, the development of pubescence, seminal fluid and menses, are not as significant in term of structural reformation as they are in the foundation of physiologic activations which tremendously affect ego reactions as manifested in social relationships. Self-consciousness, vanity, aggressiveness, submissiveness, self-pity, anxiety, social timidity and *self isolation are reflected in the personality reaction* of the child to its own physique and physiology. Consciousness of one's growth is a psychological influence more consequential than the physical fact.

It is because of the unusual features of this phase of life that we felt that a study of the attitude of the adolescent towards his body would be of interest.

We were all aware of the tremendous effect on personality that marked structural or functional defects may have. The loss of an extremity or the rheumatic heart may profoundly alter a child's attitudes towards himself and towards his environment. This problem has been considered in many studies (Forsythe, 6; Palmer, 11; McHale, 10; Barraclough, 1). However, it is not these problems that we wish to concern ourselves with here, although in passing we may touch upon them. We are more interested in the reactions of the child to certain morphologic or functional aspects of the body which on the whole fall within the limits of normality or do not deviate sufficiently from the norm to be considered markedly pathological. The personality of the individual is intimately bound up with his body. Schilder (13) expresses this explicity when he points out that:

> Bodies are after all not isolated entities. The body and the body image are always the body and body image of a personality which expresses itself in the body. The body image is never an isolated part of our existence but is a part of every experience. The human personality is a personality with a body which expresses itself in the body image and only on the basis of the understanding of the body image can we understand the personality fully.

After these preliminary observations we turn to an examination of our material in the hope that we might be able to shed a bit more light on the rôle that the body plays in the personality of the adolescent boy.

In the study 100 unselected cases were taken, in the order of admission. As can be seen from Table 1, the majority of the

TABLE 1

		Group I	Group II	Group III	Group IV
Totals		43	24	30	3
Color	W	31	13	29	3
	B	12	11	1	0
Religion	Catholic	28	14	22	
	Protestant	13	10	4	
	Hebrew	2	0	4	3
Intelligence	H. average	3			
	Average	19	11	11	3
	Dull normal	8	7	8	
	Borderline	4	2	8	
	Defective	7	4	1	
	Unclassified	2		2	
Diagnosis	Conduct dist.	37	18	9	
	Neurotic	1	3	15	
	Psychopathic personality:				
	Asocial & amoral	3		1	
	Schizoid	1	1		
	Cyclothymic		1		
	Organic brain disease		1	3	
	Undiagnosed	1		2	3

patients were delinquents with varying conduct disturbances. We designated those children as neurotic in whom we felt that the clinical picture, be it delinquency or otherwise, was a manifestation of a neurotic conflict. Before proceeding any further, we should like to elucidate upon our method of approach. We quite agree with David Levy (7) when he points out the disadvantages of the questionnaire method of approach to problems of this sort. However, for the sake of uniformity and convenience we found that the questionnaire type of approach was fairly satisfactory as it all depended on the attitude of the examiner, the ease and naturalness with which he approached the problem. The questions more or less tended to

follow those formulated by Schilder (14) in his studies of the body image. The examiner tried not to limit himself rigidly to these questions but went along with the child if the latter wished to elaborate. For purposes of clarity we are presenting the questions here in groups—although this was not the form followed in the actual study. Under the first heading we included all those questions which concerned themselves with the head and parts of the head. The second group of questions discussed the extremities. Under the third group were placed those inquiries which had to do with the torso (chest, breast, abdomen, buttocks, etc.). The fourth group of questions was centered around the genitals. The final series concerned themselves with some of the general functional aspects of the body and some morphologic aspects not included under the above headings. The individual questions will be referred to when the groups are discussed in more detail.

In evaluating the responses we were guided by the type of reply and by the interest the child showed in answering the question, as revealed by his attitude and by his elaborate or noncommittal replies. These were some of the factors which determined our evaluation of the child's responses. In the tabulation we were only concerned with active responses; that is, those which showed some reaction, be it positive or negative, which differed from the usual response. If a child was asked a question about some particular part of his body, for instance—*"What do you think of your nose?"*—and his response was noncommittal, such as— *"It's all right,"* or *"I'm satisfied,"*—it was considered as an inactive response. Only active responses either in a negative or positive sense were tabulated, if he said, for instance,—*"My nose is too long,"* or *"I have a very nice nose compared to other boys."* In other words, what we were mainly interested in were those responses which expressed some thoughts, be they positive or negative, stronger than the average run of responses. In this way we were able to separate out, rather crudely, four groups. This separation was not done statistically but was more or less based on the general impression of the examiner during the interview in evaluating the responses and the affective reactions of the child as a whole. In the first group were found the group of children who showed very little body interest, although giving active responses to many questions. This group was clinically closely allied to the second group where there were more bodily interests expressed which, however, did not appear to influence the clinical

picture. In the third group were placed those cases in whom the body interests were intimately bound up with the behavior pattern of the child although the preoccupation with the body was not the clinical picture. Finally we come to the last group in which preoccupation with the body was an outstanding feature of the clinical picture. The rather interesting picture which the cases in this last group showed and the rather limited number of cases impelled us to omit them from the general tabulation and discuss them in more detail.

It can be clearly gathered from the figures that out of 100 cases almost half fall into the first group, i.e., those in which the activity of responses appeared to indicate little body interest. Clinically, as was pointed out, there appeared to be a close relationship between the first two groups and we can readily see that the number of cases in these two groups comprised two-thirds of the cases. One-third fell into the combined groups of 3 and 4, the former having 30 cases and the latter three. Although the majority of cases fell into the first two groups it is interesting to note the large size of the third group—that in which the body interests were intimately bound up with the behavior of the child. However, in all groups, especially the first two, the inactivity of responses was quite noticeable. We cannot, of course, tell if among these inactive responses there would not have been more active ones, if the children were more honest in their responses.

Before proceeding to the analysis of the individual question groups, let us try to obtain a kaleidoscopic view of the various groups in an effort to establish some trends under the headings of color, religion, intellectual capacity, and clinical picture.

The percentage of negroes on our ward is fairly high, almost one-fourth of the children at times being negro. In a recent follow-up study of the first 300 court cases on the adolescent ward, Carrol and Curran (3) found 24.6 per cent were negroes. We notice in Table 1 the gradual diminution of negroes as we approach Group 4. These children were, on the whole, much more aggressive than the others, and the majority of them were sent in because of conduct disturbances. Occasionally we came across a colored child somewhat given towards introspection and in him the picture usually showed some rather unusual features, as in the case of *C. L.*, a defective colored child who apparently suffered from a psychotic episode following a head trauma. This child as he recovered

showed some preoccupation with the "bumps" on his face, meaning by that the pimples. He would walk around on the ward rather preoccupied with this and whenever he had an opportunity would stop and ask the physician to do something about the "bumps on my face." On the whole, the attitude of most colored children was one of surprise at the questions asked as though "a person must be crazy to ask questions about such perfectly natural things." There was a general acceptance of the body without explicit interest. This was more or less characteristic of the children in Group 1. The questions on sexual matters were handled much more naturally by most colored children with very rare exceptions.

Regarding religion, very little of significance was brought out beyond the fact that Jewish children appeared with somewhat more frequency among the groups with more body interests—Groups 3 and 4. This more or less coincides with the popular conception that the introspective type of individual is frequently to be found among Jews. In justice to the problem, however, it might be pointed out that the majority of the Jewish children admitted to our ward (at most always a decided minority) were essentially not conduct disturbances. It is, therefore, not permissible to attempt to draw conclusions from these observations on to the general population. In a personal communication Schilder suggests a possible explanation for the finding of Jews among the groups with more body interests. He points out that insecurity of any kind leads to a difficulty in integrating the parts of the body into a unified whole. The consequence of this difficulty leads to a greater interest in the individual parts of the body rather than in the body as a unit.

In examining the distribution of cases under the heading of intellectual capacities one is struck by the relatively high number of patients with limited intelligence seen in Group 3. The nature of our material is such that we have quite a high distribution of children with limited intelligence. However, relatively speaking, a fairly high percentage was found in Group 3. The exact significance of this is not quite clear. We did note, however, that frequently children with limited intelligence, who were dimly aware of their inadequacy in coping with the environment, would stress certain minor defects of their bodies in an effort to express this inadequacy. Here again Schilder suggests that individuals with less intelligence might have no reason to repress defects if these were present, or they might not be able to integrate these defects into the body as a whole. This is a gen-

eralization whose justification would require deeper study of the cases than we were able to undertake.

The predominance of conduct disturbances in all groups is to be expected from the nature of the material which is admitted to our wards. However, it cannot fail to strike one that as we approach Group 4 the percentage of personality variances increases and we find more personalities who deviate from the norm. Out of 33 cases in Groups 3 and 4, 24 showed clinical features which took them out of the usual run of conduct disturbances, such as psychopathic personalities, organic brain diseases, or neurotic reaction patterns. The explanation might be given by Schilder's (13) indication that where we have disturbances in libidinal relationship, or where insecurity comes into play, there is also a disturbance of the body image with a destruction of the integrative power of the personality with respect to the body image. We might here again mention a word about the diagnosis of Conduct Disturbance and Neurotic Reaction Patterns as almost one-half of our cases in Group 3 fell into this category. Conduct Disturbance was essentially a descriptive diagnosis of the offense for which the child was admitted and was usually made where the dynamics of the case were such as to point away from a neurotic conflict. Under Neurotic Reaction Pattern we did not necessarily include outspoken neuroses but rather those cases, who although descriptively demonstrating similar behavior to that of the conduct disturbance, i.e., truancy, stealing, etc., did so as manifestations of a neurotic conflict.

We may now turn to an examination of the responses to the questions proper. It is not the purpose of this paper to make a rigid statistical study as there are too many variables to be contended with. We do feel, however, that several trends might be indicated which if properly evaluated might be pertinent. In a comparison of the average total responses in all the groups including 3 and 4 together, we note that there were twice as many responses in these two groups as there were in the first two groups. The actual figures do not clearly bring out the emotional reactions of the children during the interview. There was very frequently a good clue as to the amount of interest the child had in his body and undoubtedly this was much more pronounced in the last two groups.

Let us now turn to the type of responses to questions on the parts of the head. Here were included questions on shape of the head, eyes, nose, mouth and teeth, skin and complexion, hair, ears, etc. The

question was worded usually as—"*What do you think of your eyes?*" or "*What do you think of your nose?*" etc., and was elaborated upon if the need was felt to do so. A quick glance reveals that the negative type of response, those of dissatisfaction, far outweigh those of satisfaction and the appearance, rather than the function, appears to be the factor stressed in all groups. In reviewing the individual elements one can see that the nose, mouth, and teeth appear to be emphasized in all categories. The response to the questions on mouth and teeth were almost all given in connection with the teeth and concerned themselves mainly with the cavities, which were described from a morphologic rather than a functional point of view. This was more or less conditioned by the fact that all children who were admitted to the ward were routinely given dental care and made aware of the treatment necessary.

The nose comes in for a goodly share of preoccupation. This is not the place to go into a detailed discussion on the significance of the nose to the general physiognomy. However, we might indicate to what a marked degree this portion of the anatomy is stressed socially. The value placed by society on a small nose, rather than a large misshapen one, as a racial characteristic can be seen in the discussion of our adolescents. This is noted in all groups but is somewhat more outspoken in Group 3, and as will be seen later, in the cases of Group 4. Of course, one must not ignore the fact that a large misshapen nose, quite objectively, does distort the looks of a person, especially if it is out of proportion to the rest of the face.

Turning to responses to the questions on complexion, it is not at all surprising that this part of the face comes in for a goodly share of comment when one considers that with the advent of puberty acneiform complexions make themselves manifest. This was looked upon as a decided disadvantage to some of the older children who lay so much relative stress on their facial appearance. One frequently heard expressed the popular conception that "pimples" came from excessive masturbation and from stomach disorders, although the majority of the children attributed their poor complexion to the fact that they did not have individual soap and were thereby contaminated. To a certain extent the preoccupation with complexion was somewhat conditioned by the fact that many of the more pronounced cases received ultra-violet radiation so that many of the children, who may not have paid much attention to their complexions under ordinary circumstances, did so after admission

to the ward. The most pronounced preoccupation with "pimples" was found in some of the cases in Group 4 which will be discussed later on. Color might be mentioned here, as it undoubtedly must play a big rôle in the life of the negro child. In brief, our examinations revealed that a frequent type of response received from the negro child was a defensive one, worded either, *"I'm satisfied with it"* or *"That's the way God made me and I'm satisfied."* A general trend made itself apparent to value a lighter complexion more highly than a darker one. That this is a socially conditioned type of response is quite apparent and needs no further discussion here.

Turning now to a discussion on the responses to the question on the extremities, it is interesting to speculate why there were more active responses in the last two groups than in the first two. In the main it was found that the children in the last two groups were clinically, on a conscious level at any rate, inclined to be more passive. They were the ones who on the ward were the victims of the more aggressive children and felt their inadequacies in the neuromuscular apparatus quite keenly, this being one of the main means of competitive expression in puberty. The more aggressive children and those who were somewhat older apparently accepted their extremities as well functioning on the whole. There were many more expressions of satisfaction with the extremities by these children than in the children who were found in the 3rd and 4th groups, where only one response out of about 78 was of a positive nature. One of our cases was that of a 16-year-old boy who demonstrated extreme aggressive drives against other children. These were expressed in incidents such as throwing acid in another child's face, pushing a child into a fire, strangling boys with his hand, and putting poison in a pot of food. This boy pointed to his extremities as the most important parts of his body, insisting that his hands were more important than his heart or brains. Many of the children in the 3rd and 4th groups, who were frequently victims of aggressive attacks by the other children, complained bitterly of their ineffectual extremities, ineffectual not only in defense but also in competitive sports. This deficiency in the neuromuscular apparatus apparently takes on more importance with the advent of puberty. Many of the children were not aware of this inadequacy until they reached adolescence. This is the age when aggressive competition is stressed in the neuromuscular apparatus, not only in play on the streets but also in the schools. It might be pointed out that the

nature of the cases on our ward was such as to have these particular phases of competition stressed rather than intellectual ones. This does not mean that the neuromuscular apparatus does not play a rôle in earlier years, but certainly not to the extent that it does with the advent of adolescence. In line with our discussion, one would expect to see the functional rather than the morphologic aspects stressed. If the inadequacy is felt in the sphere of neuromuscular expressions of aggression or defense against aggression one would expect to find this expressed in a functional rather than morphologic sense. However, one is surprised to see that the opposite is true and is at a loss to explain this discrepancy. This becomes clarified when it is pointed out that although the expression of the response was morphologic the actual intent behind it was distinctly functional. To illustrate: A response to what a child thought of his arms when given as—*"They're too skinny and thin"*—is superficially morphologic. However, it was apparent to us, from the general picture, that this was meant too skinny and thin for competitive purposes in a fuctional sense. When the child said his hand was too small, it meant much more than merely it looked too small, but was not large enough for defense or offense. If one looks upon the responses in this light the preponderance of functional responses over morphologic makes itself quite apparent. The type of response is here again overwhelmingly one of dissatisfaction, just as we found it in our first group of questions.

The parts of the body included under torso were the chest, the nipples, abdomen, and buttocks.

Our investigation revealed that the preoccupation with the torso is somewhat less than the previously discussed parts of the body. From the nature of the material, it is apparent that the responses are expressed in a morphologic sense. Again we note that where there was a response it was inclined to be in a negative sense; only in Group 1 did we see some positive responses, practically all related to the chest. Although here again the response is made morphologically in many cases it had a functional sense as when a child will point out the fact that his chest is big, the significance being that it can take punishment.

The questions on the genitals brought a paucity of responses, the general attitude being one of nonchalance and acceptance. However, one interesting observation stood out. This was a rather general tendency to belittle the size of the penis in the cases where it was

large or of normal size. It is quite apparent that a large penis would not have the significance for our patients—the majority of whom have not indulged in heterosexual relations—as for adults in whom we all know the evaluation of the size of the penis plays a big rôle. However, there was a further reason for this tendency to belittle the size of the penis and say it was smaller than it actually was. This was expressed by many children who pointed out that one got a large penis from masturbating too much and a boy with a large penis was frequently accused of this practice by others. A goodly number of our cases actually did have small genitals and in those cases where the hypogenitalism was pronounced it was pointed out by the child.

In some cases the question as to the possible existence of a bone in the penis was asked. A goodly portion of the responses indicated the belief that there was a bone in the penis and only in two cases was the proper response of the vascular origin of the erection given.

We now turn to a discussion of the questions on general body functions. These concern themselves with questions on strength, health, lungs, digestive tract, and urination. It is interesting to note the high number of responses in Group 3 given in answer to the question—"*What do you think of your strength?*" Out of the 30 cases there were 22 responses of which 19 were expressions of dissatisfaction—a feeling of weakness in comparison to other children. This trend is more or less to be expected from the nature of the material in this group as was pointed out above. This tendency was noticed in all the questions of this group and, relatively speaking, the activity of responses was quite markedly increased in Group 3 over the others. An interesting observation is noted when one examines the responses to the questions on breathing. Here we received active responses, all in a negative sense, in a goodly number of cases. A complaint met with frequently in our children was some form of difficulty in breathing, expressed often as—"*Sometimes my breathing is hard and gets stuck.*" Perhaps we are dealing in these cases with mild anxiety equivalents. We do not know. It is interesting to note that gastro-intestinal complaints were few. This is a finding which differs somewhat from hypochondriasis where *G. I.* complaints play a fairly big rôle. Our findings seem to indicate that the cases of body preoccupation which are described in this paper are not essentially of a hypochondriacal nature.

Still concerning ourselves with the functional aspects of the body,

we sought responses to questions as to the best or worst functioning parts of the body. To further bear out the importance of the neuromuscular apparatus to the adolescent, it is interesting to note that in all groups the extremities were stressed in response to these questions. There was, however, this difference in that Groups 1 and 2 tended to choose the extremities as the best functioning parts of the body and Groups 3 and 4 tended to choose them as the worst functioning. This is an observation which coincides with our above findings. However, the outstanding feature in this group of questions was the great number of vague and inconsistent responses in all groups. The patient appeared to have difficulty in expressing himself clearly as to the best and worst functioning parts of the body. Schilder (personal communication) feels that this vague and confused attitude expressed the desire of the child to integrate the parts of the body into a unified whole and a reluctance to consider the parts in themselves. This difficulty in response is more noticeable especially in the replies of Groups 3 and 4 to the question as to the worst functioning part of the body. This seemed to express a general attitude of refusal to admit the poor functioning of any part of the body. A frequent response was: "*No part works poorly. Everything is all right.*" This scotoma sometimes assumed large proportions as was seen in the case of one child whose right forearm was missing and who, at the time of examination, was just recovering from an attack of rheumatic fever with cardiac involvement. This patient completely ignored the deficiencies of the heart and arm and only when his attention was called to them did he admit, rather reluctantly, their poor functioning, adding immediately that as far as his heart was concerned he was already recovering and in a short while nothing would be noticed. Another child, who had an almost fixated finger as the result of an old injury, throughout the whole questionnaire ignored this defect. When his attention was called to it he admitted that he was somewhat limited but in an obviously over-compensating fashion blithely spoke of the future correction of this hopeless defect. Whereas this type of reaction was true in all groups it was more pronounced in the more aggressive children who in many ways were obviously over-compensating. This attempt at overlooking an obvious defect was done quite consciously and differs, therefore, from the anosognosias of Babinski in which, according to Schilder (13), we are dealing with an organic repression. Here again the unsatisfactory type of response stands out in

the questions as to the nicest and ugliest looking parts of the body. Several factors were responsible for this among which that of modesty played a big rôle. The question was frequently evaded with a rather embarrassed laugh—*"Well, all parts are nice"* or a similar response of that sort. Here again we might point out as Schilder does, the tendency to integrate the parts of the body into a whole, as a possible explanation for this attitude on the part of the children.

The fact that the head and face are stressed is quite natural, that being the portion of the body in which the factor of looks plays such a big rôle. An interesting sidelight is seen in two cases where the shape is pointed to as the nicest part of the body. Both these cases who belonged to Group 3 were dysplastic, obese types of children and in one the shape was undoubtedly to a great extent contributory to his difficulties in adjusting as the boys used to ridicule him because of its somewhat feminine outlines and obesity.

The question as to the ugliest part of the body gave the greatest number of responses to the genitals and buttocks in all groups. Although this was more or less expected, the vehemence with which this was asserted in some cases was surprising. It took the shape of castration wishes in one case who hated his genitals. This child rejected sexuality completely to the point of wishing for an asexual life in the future. Another case of this sort was seen in a colored child. This was somewhat unusual as the colored children more or less accepted sexuality more naturally than did the white.

Under this group of questions we also touched on the future bodily development of the child. One quite readily noticed here the rather constant desire for increase in height and strength as a future developmental wish. A few cases in prognosticating as to their future development, in a rather resigned fashion, pointed out that they would probably be short, and some that they would be weak. The usual reasons were given as familial—*"My father is that way,"* etc. The desire for height and strength was quite definitely an expression of wish for superiority in competition. When asked why they wished to be tall in almost all the cases the child would say that he would then be able to handle others who might be too aggressive. This trend made itself apparent in most parts of the body which were used in competitive situations.

We have thus far considered the cases in the first three groups. We will now describe in some detail several cases in which body preoccupations were outstanding features of the clinical picture.

The first case was that of a 16-year-old Hebrew boy (*M. W.*) who was brought into Bellevue by the parents at the suggestion of the Adjustment Bureau of the Domestic Relations Court. According to the information we received this boy appeared at the court asking to be sent to prison as he was no good, a vagrant who should be punished by imprisonment.

The physical examination was essentially negative. Psychometric revealed him to have an *IQ* of 87 on the Bellevue Intelligence Test which would make him of Dull Normal Intelligence. However, it was the impression of the psychologist that the boy could do much better and was probably of high average intelligence although he was not functioning at his full capacity in view of his disturbance.

In appearance the boy was not bad looking although he had a rather slovenly, slouching carriage. His hair was unkempt and unbarbered. He had a somewhat acneiform complexion and slightly hooked nose, which, however, was not pronounced.

On the ward the boy was at first somewhat quiet, seclusive, and preoccupied. He preferred to remain off by himself and partook of the ward routine only reluctantly. As time went on he became somewhat more talkative and friendly with the other children so that finally he became fairly well liked by them although they looked upon him as somewhat queer and they continued to tease him about his appearance. During an interview he revealed marked feelings of inferiority and inadequacy, many of his preoccupations centering around his looks and his appearance. Because of this he was constantly being criticized by his family and associates. As the child put it:

> They called me a piece of meat and a pair of eyes. My sister picked on my appearance and said I looked and walked like an old man. She would always say I didn't look good enough. My father, too, was always joking and making fun of how I looked. I think they're right.

The members of his family admitted that they would constantly call attention to his unkempt hair, to his slovenly appearance and poor carriage. Finally, in a mood of complete humiliation and rejection he ran away—*"I thought I had so many faults I was a burden to my family. I was no good, so I ran away."*

The parents were not the only offenders, as throughout his years in school the boy's appearance was subjected to criticism and wit-

ticism by both teachers and schoolmates. The other children would call him "Speedy" or "Creeping Jesus" because of his rather slow way of doing things. It is interesting that this boy's breaking point took place at puberty, this being somewhat in line with our introductory remarks as to the importance of the body image to the child at puberty.

Throughout the questionnaire one found trends pointing to the above picture. We will not report the whole protocol but merely reproduce some of his characteristic responses.

About his nose the boy says the following: "I know I have a big nose; I don't like its shape. I don't like the tip of my nose. It's not straight. It's not like other noses."

Referring to his complexion he says: "I used to have a lot of pimples but that treatment (ultra violet) cleared it up. Pimples ruin your appearance. I used to feel very bad when my face was full of pimples, I felt inferior."

When asked what he thought about his body in general the boy said: "I have an inferiority complex about it—when I compare it to other boys I feel inferior about it."

The picture, therefore, is that of a boy whose appearance was subjected to constant criticism both by the family and other members of the community. He developed marked feelings of inferiority and inadequacy leading to an episode characterized by complete humiliation and surrender. He shows as a major part of the clinical picture marked bodily preoccupation which was reflected in the questionnaire.

Turning to the next case we have the problem of a 15-year-old Hebrew boy who was brought in by the mother at the suggestion of the school authorities as the boy had refused to go to school for the past six months. He would not go out of the house, preferring to stay at home in bed. He expressed the idea that his face was disgusting, his nose was too long, and that he had too many pimples on his face. While at home he would constantly exercise, drink fluids and eat raw vegetables in an attempt to clear his face of pimples. He would not touch sweets or any foods which he felt would contribute to the development of pimples. He would closely watch the effects of this régime on his complexion and would constantly look in the mirror.

On admission it was seen that the ideas expressed by the boy had their basis in fact in that he did have a large nose and a markedly

acneiform complexion. He quite readily admitted the facts given in the above history adding that the boys at school used to make fun of his appearance and he was extremely sensitive about his nose, feeling that its size was quite noticeable. He showed some preoccupation with sexuality but pointed out that because of his appearance, his desires for female companions could not be satisfied. "I was just not satisfied with my looks. I'd look in the mirror and I'd become unsatisfied—I didn't like my nose, it was too big. The whole thing started with those pimples on my face—they were disgusting."

Throughout the boy's stay at the hospital, his interests continued to be centered on his body. He would frequently request to be sent up to the roof so that he could exercise and be healthy—which would eventually lead to the clearing up of his skin. He became somewhat intrigued with the idea of having a plastic performed on his nose and made plans to have this procedure at a future date.

We see here a case, therefore, where the clinical picture is predominated by an intense preoccupation with facial features—a preoccupation characterized by marked feelings of inferiority and inadequacy brought out by the repeated criticism of his environment.

Another case (*H. K.*) is the problem of a child whose constitutional inferiority was intimately bound up with his inability to adjust in his interpersonal relationships. The full account of this boy's background is too detailed to be given here. Suffice it to say that this boy was unable to get along with children practically all his life. They would make fun of his appearance calling him "Flapjack" and "Clark Gable" because of his large ears. He was somewhat dysplastic and showed poor neuro-muscular coördination so that he was unable to indulge in competitive sports successfully. This also would be a cause for subjecting him to the witticism of boys his own age so that he resorted to playing with children younger than himself. We have seen this reaction quite frequently in some of the smaller and weaker children on our ward. In reaction to this constant abuse and criticism, the boy developed a system of delusion which led some observers to believe that he was schizophrenic.

In appearance, as was mentioned above, he was somewhat disproportionate. His head was slightly hydrocephalic and he had large ears which stood out almost at right angles to his head. These were the main features which were chosen by the children on the ward as the focal points of their witticism. They, too, christened

him "Flapjack" and "Clark Gable" among other names and the child's life was made miserable as a result of his treatment by the other children.

The questionnaire demonstrated the child's reactions to this treatment. When asked what he thought of his strength and appearance, he said:

> I'm weaker than the average. Other fellows can push me around and if I push back they end up the strongest and I cry and walk away. If I played basketball, I'd end up in the hospital with bruises or concussions of the brain. . . . I used to think I was ugly with my big ears and my face but I think I'm getting more handsome as I grow older. My ears are very big. I have a big head and a skinny body. I think I act like a sissy. My face is handsome like a girl's face. My face is the wrong shape. I'd like it pointed so it would look like a skull. . . . My arms are too skinny. . . . I want to be taller, full grown like my father.

There are many more features in this case which, although interesting cannot be given here. He illustrated, however, the forceful effects that group criticism of the body may have upon a child.

In this group it is quite evident that the clinical picture is markedly predominated by an excessive interest in the body. It could quite justly be pointed out that this was only one feature in a process which involved the total personality of the individual. However, the expression of the conflict in terms of body image is what is so outstanding in these cases. These children did show, objectively, some deviations in body structure; however, this does not explain why the preoccupation should have assumed such marked proportions as to be the outstanding feature in the clinical picture. It must be made quite clear that we are not dealing with hypochondriasis in these cases. We feel that the process is a different one. This is not the place to go into a discussion of this problem, suffice it to say that other authorities (Schilder, 13) who have studied the body image thoroughly have a similar viewpoint.

The tendency is for the normal person to view his body as a whole. Wile (16) points out that—"The entire nature of the child is represented in his total reaction as an organism. He is not an assemblage of organs etc." What happens, therefore, to the unity of the body in the cases of this group? Why is this unity so disturbed that parts begin to stand out, to demand the patient's atten-

tion? Schilder (12) points out: "The unity of the body image is disrupted when libido is too unequally distributed over the body and when some parts are invested with too great an amount of libido." He adds that emotional influence will change the relative value of the different parts of the body image according to the libidinous tendencies. This process becomes pronounced as we approach psychoses, where the unity of the body is destroyed. In a paper on alcoholic hallucinosis, Bromberg and Schilder (2) discuss the dismembering motive in which the whole unity and integrity of the body is destroyed. Whereas there was a question as to whether the cases we described were psychotic or not, it is certainly apparent that the unity of the body image has in some way been disrupted.

Discussion

The individual features of our study were discussed in some detail as we took them up. We prefer at this point to go into some of the aspects which we feel are worthy of note. We indicated above that what we included in the tabulation were active responses, those in which the subject, either in a positive or negative sense, indicated some deviation from the usual noncommittal response. The latter type of response was the most frequent in most of the cases. The normal children tended to look upon their bodies as a unified whole, and it was extremely difficult for them to comprehend the examiner's questioning upon individual parts. There was a disinclination in the majority of the cases to break up this unified whole and view it in parts. In those cases where the child gave excessive attention to the parts of the body and where there was difficulty in integrating them into a unified whole, we were dealing with cases who showed personality distortions. One could say that there was a fairly high correlation between social adaptability and interest in the body. When the body is experienced as a whole then we were dealing with children who did not reveal personality distortions, whereas those cases who showed excessive body interests revealed personality distortions with difficulties in social adaptability. It is, therefore, apparent that no clinical or analytical investigation is complete before one knows about the body image and the ideologies attached to it.

If any particular outstanding impression is made upon us as we study our material and review the attitudes of the children during the examination, it is the powerful influence of the group upon the

child's attitude not only towards his own body but towards the bodies of others. In his own study Levy (8) points out the outstanding influence that parents have on the child's attitude towards his own body. However, even he saw the group as a moulder of the child's attitude. He points out that:

> The individual's response to discrepancies between his anatomy and his concept of the group norm is also reinforced by the fact that the group picks out the individual's discrepancies for him and makes them the target of criticism, approval, and disapproval. When the group gives special reassurance or approbation it adds to the individual's self-assurance, when it ridicules, self-esteem is lowered.

This is very apparent in our study where the influence of the group brings itself to bear on the child's attitude towards the body in the sphere of competition. The use of the body, competitively, appears to make itself quite evident in puberty and this was brought out clearly in our study.

How the group may influence the attitude of a child towards his body was brought out when an attempt was made to interview the children on their attitudes, in groups. It was then apparent that several children were guided in their responses by the remarks of others. These same children when seen alone admitted they were influenced by others in their responses and displayed a different attitude.

In his book on the body image Schilder (13) devotes a whole chapter to the question of the "Sociology of the Body-Image." He points out that the social relation is not only a relation between two personalities but is also a relation between two bodies. The conclusions which Schilder arrives at in his chapter are quite well corroborated by many of the findings in our study. He points out that the attitude towards the different parts of the body can be determined by the interest persons around us give to our body. But this is not all; the interests others have in their body will influence the interest in the respective parts of the subject's own body. This too was clearly brought out in many of our cases. We saw children who did not have acne, begin to pay more attention to their complexions because some of the older children were being treated for acne.

Schilder (13) raises the question of identification and the enormous part that this mechanism plays in the building up of the

body image. "The body image of others and their parts can be incorporated completely into our own body image and can form a unit or they can simply be added to our own body image and then merely form a sum." It was more imitation than identification which we found at play in our study. However, we know that imitation frequently leads to identification. Thompson (15) in a recent study examined cases in which there was an identification with the enemy. She pointed out that many children used this as a method of mastering the environment by identifying themselves with the aggressor. The child for protective purposes will join forces with a hostile power. This was a reaction frequently seen in the children on our ward where many of the more passive children would after a while become quite aggressive and in turn pick on other children more passive than themselves (5). During interviews it was quite easily brought out that this was purely a defensive measure resorted to by joining forces with the more aggressive children. Although this was frequently quite conscious, Thompson points out that—"A conscious tendency to go through the motions of being like other people is frequently found in individuals who also have a tendency to make identifications." This tendency to identify with the more aggressive children made itself also manifest in the attitude of the child towards his body. Many of the children who heretofore were more concerned with different parts of the body, parts not included in the neuromuscular apparatus, began to show some interest in these parts. They began to watch the growth of muscles in their arms and to point out any progress to the examiner. This change in the attitude of the child towards the organs which are used for the expression of aggressive competitive drive is undoubtedly most pronounced at puberty. On our ward children who had never put on a pair of boxing gloves before began to take cognizance of their extremities in terms of "reach"—"footwork"—etc. Levy (8) points out that "the wish to grow up, to be big, is developed in children through training, imitation, or rivalry with the parents, competition among siblings or companions." However, he goes on to point out that "playmates have a much greater effect in instilling and strengthening such values than do parents and teachers. These values are corrective of parent-child dependencies."

The child's concept of his body is bound to change as he enters adolescence because the activities which he indulges in now are of a different nature than they were previously. Schilder (12) points

out the changeability of the body image and says that "the process of building up the body image is one of continual active development and through action and contact with the environment this assumes different forms."

The children on our ward were quite relentless critics of minor discrepancies in the bodies of other children. Nicknames were usually built up around these discrepancies, appellations such as "Shorty" or "Fat" were quite common. However, even more rare abnormalities were discovered and utilized by these children as was seen in the case of one boy who had ichthyosis and was called "Elephant Skin" or another child who although 16 years of age had no pubic hair and was as a consequence called "Hairless Joe." These children were quite sensitive about the nicknames, and during interviews although pretending to disregard them, they reluctantly admitted that they would prefer not to have these nicknames.

The influence of the group and society on the attitude towards the body is excellently demonstrated in some of the totalitarian countries where interest in race hygiene and anthropology has taken on such large proportions that it is not at all uncommon to hear lay people express opinions about their own and other people's body formations. Terms like "Aryan" and "non-Aryan" with their anthropologic connotations have become part of the philosophical concepts of these countries and have crossed the boundaries into other lands. It is inevitable that such an awakened interest in bodies and body image concepts should bring about a change in the individual's attitude towards his own body and towards the bodies of others. Certain features which heretofore were not bound up with much affect or interest may begin to be looked upon as decidedly undesirable and vice versa.

We hope we have been able through the above discussion to demonstrate the important rôle the group plays during adolescence in the development of the child's attitude to his body. We feel that during this period of his life the group is one of the outstanding factors at play and is one which we have to deal with in all our contacts with adolescent boys. We truly agree with Schilder when he points out that:

> Our own body image gets its possibilities and existence only because our body is not isolated. A body is necessarily a body among other bodies. We must have others about us.

Summary

In this paper a study is made of the attitude that adolescent boys have towards their body and parts of their bodies. It was felt that the changes during adolescence brought with them a different attitude of the child towards his body. The method of approach was essentially the questionnaire which, however, was not strictly adhered to. The questions were grouped under several headings such as head, extremities, torso, genitals, and general functional and morphological aspects of the body. Only active responses were tabulated, active either in a positive or negative sense. It was found that, in the main, the responses were inactive, indifferent, and noncommittal. However, in the classification of active responses it was found that the cases crudely fell into four large groups according to the extent of body interest manifested. In the first group were children who showed very little body interest. In the second there were many more active responses expressing body interests which, however, did not appear to influence the clinical picture. In the third group were those cases in whom the body interests were intimately bound up with the behavior pattern of the child although the preoccupation with the body was not the clinical picture. Finally, there was the group in which preoccupation with the body was the outstanding feature of the clincial picture.

An attempt was made to correlate this grouping with various factors such as color, intelligence, clinical picture, etc. In the main it was found that negro children appeared most frequently in Groups 1 and 2 and suggestions as to why this was so were discussed. It was felt on the basis of division of cases that children with more limited intelligence might have some difficulty in the integration of the body image, as a fairly large number, relatively speaking, were found in Group 3. The majority of the cases in Groups 3 and 4 differed markedly in their clinical picture from those found in Groups 1 and 2 who were in the main conduct disturbances, presenting few features in the clinical picture which deviated from the norm. Reasons for this were discussed in detail. The relatively greater interest in the neuro-muscular apparatus expressed by children in Groups 3 and 4 were discussed and a possible explanation postulated for the dissatisfaction with these organs in the competitive sphere. Almost all of the responses were expressed in a negative sense of dissatisfaction although some positive responses were found in Groups 1 and 2. There was a tendency to underestimate the

size of the genitals, as a large penis was looked upon as an indicator of excessive masturbation.

In the main the responses indicated a disinclination on the part of the majority of the children to disrupt their body images and view individual parts. In those cases where there was an excessive interest expressed there was found that a disruption of the body image was associated with personality disturbances. Finally, the influence of the group and society upon the body image concepts in the adolescent was discussed as one of the outstanding contributions of this study.

References

1. Barraclough, W. W. Mental reactions of normal children to physical illness. *Amer. J. Psychiat.*, 1937, **93**, 863.
2. Bromberg, W., & Schilder, P. Psychologic considerations in alcoholic hallucinoses: Castration and dismembering motives. *Int. J. Psychoanal.*, 1933, **14**, 206.
3. Carroll, H., & Curran, F. J. Follow up study of three hundred court cases from the adolescent ward of Bellevue Hospital. *Ment. Hyg.* (In press.)
4. Curran, F. J. Organization of a ward for adolescents in Bellevue Psychiatric Hospital. *Amer. J. Psychiat.*, 1939, **95**, 1365-1386.
5. Curran, F. J., & Jenkins, R. L. Evolution and persistence of groups in a psychiatric observation ward for adolescent boys. *J. Genet. Psychol.* (In press.)
6. Forsythe, D. Psychological effects of bodily illness in children. *Lancet*, 1934, **2**, 15.
7. Levy, D. A method of integrating physical and psychiatric examination. *Amer. J. Psychiat.*, 1929, **9**, 121.
8. ———. Body interests in children and hypochondriasis. *Amer. J. of Psychiat.*, 1932, **2**, 295.
9. Lhermitte, J. L'image de notre corps. Edition de la Nouvelle Revue Critique, Paris, 1939.
10. McHale, K. Comparative Psychology and Hygiene of the Overweight Child. N. Y. City Bureau of Publications, Teach. Coll., Columbia Univ., 1926.
11. Palmer, H. D. The organic aspects of behavior disorders in childhood. *Penn. Med. J.*, 1937, **41**, 186-191.
12. Schilder, P. The somato-psyche in psychiatry and social psychology. *J. of Abn. & Soc. Psychol.*, 1934, **29**, No. 3.
13. ———. Image and appearance of the human body. *Psyche Monog.*, No. 4. Kegan Paul, 1935.

14. ———. Psychotherapy. New York: Norton, 1938.

15. Thompson, C. Identification with the enemy and the loss of the sense of self. *Psychoanal. Quar.*, 1940, **9**, 37.

16. Wile, I. Physical and physiological growth as a factor in child adjustment. *N. Y. State J. Med.*, 1936, **36**, [illegible]

Bellevue Hospital
First Avenue & 30th Street
New York City

The Journal of Genetic Psychology, 1942, 60, 61-70.

IQ CHANGES IN OLDER-AGE CHILDREN PLACED FOR FOSTER HOME CARE*

The Psychological Clinic, Indiana University

JAMES W. LAYMAN

This study is the continuation of a previous one (1) but based on an increased *N* and a number of cases to whom a third psychometric has been administered. The present interest is limited to the *IQ* changes and an analysis of the clinical factors which underlie these changes. The data are based on psychometric results obtained from 120 subjects, each of whom was tested two or more times. The *N* for those tested on the same forms is 105 (1916 Revision of the Stanford-Binet, 52; Form *M*, 28; Form *L*, 25) and 15 for those originally tested on Forms *L* or *M* and rechecked on the other. Twenty-one of the subjects were tested twice on the 1916 Revision and retested on Form *L* later either as a routine recheck or where the results of the later 1916 Revision were questioned. However, no cases are duplicated in computing either the averages or the correlations for the various tests. Cases with known reading defects or where there were questions as to coöperation, test rapport, etc., were not included in the data.

As indicated in the previous study, most of the children come from families whose natural homes are of low socio-economic status and with backgrounds characterized by high incidence of parental inadequacies. The average age at placement was 135.05±33.65 months with a range of approximately 6 to 16 years. The period of foster care ranged from 7 to 166 months with an average of 25.92 months. In most instances, foster homes are located either on farms or in small towns fairly adjacent to Des Moines from where supervision is maintained by the two placement agencies.[1]

*Received in the Editorial Office on September 2, 1940.

[1]One hundred eighteen children are under the supervision of the Iowa Children's Home Society and two under the Des Moines Catholic Charities.

At this point the writer wishes to express his appreciation to the Section of Psychological Services, Iowa State Department of Social Welfare, for data on subjects examined by them.

A. Psychometric Data

Inspection of Table 1 shows that the range in *IQ* tends to increase

TABLE 1
IQ Range for Each Test

Test	*N*	First	Second	Third
1916 Rev.	52	74-117	77-123	
Form *M*	28	78-138	78-146	
Form *L*	25	74-111	72-123	
1916 Rev. and *L*	21	82-120	72-125	
1916 Rev.	22	80-117	82-122	65-134

from one test period to another, with differences between the various ranges of from 2 *IQ* points to 15 for retests on the same form. The same tendency for an increase is also found in the standard deviations of the distribution (Table 2) with the exception of the group

TABLE 2
Averages and Standard Deviations for Each Test

Test	First	*SD*	Second	*SD*	Third	*SD*
1916 Rev.	95.808	9.922	98.384	11.361		
Form *M*	101.793	14.034	106.356	15.780		
Form *L*	94.780	9.276	96.360	12.087		
1916 Rev. and *L*	97.001	11.160	100.287	14.595		
1916 Rev.	97.592	12.111	99.408	10.827	101.396	11.925

to which the 1916 Revision was administered on three different occasions. In this group, the standard deviation of the distribution is smaller on the second test but tends to increase again on the third.

The averages show a tendency to increase from one test period to another, although the differences are not statistically reliable except for that obtained between the first and second testing on Form *M* (Table 3). There are 99.4 chances in 100 of a true difference be-

TABLE 3
$Diff./SD_{diff.}$ between the Means Computed on the Formula for Correlated Measures

Test	1 and 2	1 and 3	2 and 3
1916 Rev.	2.53		
Form *M*	4.11		
Form *L*	1.23		
1916 Rev. and *L*	1.77		
1916 Rev.	.84	1.86	1.59

tween the means for the larger group tested on the 1916 Revision. In no instance is the difference between the standard deviations completely reliable, although that between the first and second test on Form *L* and that between those originally tested on the 1916 Revision and retested on Form *L* approach complete reliability. There is, however, a consistent tendency for the standard deviations to increase somewhat.

TABLE 4

Diff./SD$_{diff.}$ BETWEEN THE SIGMAS OF THE DISTRIBUTION COMPUTED ON THE FORMULA FOR CORRELATED MEASURES

Test	1 and 2	1 and 3	2 and 3
1916 Rev.	1.53		
Form *M*	1.61		
Form *L*	2.38		
1916 Rev. and *L*	2.01		
1916 Rev.	.61	.91	.57

Correlation coefficients between the results for each administration of the same form as well as that between the 1916 Revision and Form *L* are all positive and within the ranges usually reported in the literature. The coefficient between the 1916 Revision and Form *L* is somewhat lower than the one reported by Terman and Merrill (4); however, the reliability of the difference was not computed since the groups may not be comparable. The *PE* of the difference between the first and third correlation on the 1916 Revision is 2.79.

TABLE 5

COEFFICIENTS OF CORRELATION BETWEEN THE TESTS

Test	1 and 2	*PE*	1 and 3	*PE*	2 and 3	*PE*
1916 Rev.	+.775	.037				
Form *M*	+.930	.015				
Form *L*	+.852	.040				
1916 Rev. and *L*	+.813	.028				
1916 Rev.	+.621	.088	+.665	.080	+.872	.036

Neither the averages nor the correlations were computed for those initially tested on either Forms *L* or *M* since the group was too small (initial Form *M*, 10; Form *L*, 5). Differences in *IQ* between the two tests range from 1 to 16 points.

For the groups as a whole the results suggest a tendency for the average *IQ* to increase slightly from one administration to another.

The same is true also of the standard deviations of the distributions. Except for the difference between the means on Form *M* retests, the differences are not statistically reliable. The fact, however, that both the means and the standard deviations increase would indicate that the tendencies for an increase are present and that the total gains exceed the total losses.

B. Extent and Direction of Individual Changes

We have been unable to find that Terman reports any test retest correlations on the 1916 Revision. However, he (4) does state that the average percentage of change in *IQ* on a retest study was 4 per cent and that greatest difference was only 8 per cent. Table 6 shows that average percentage of change in the present group

TABLE 6
Percentages of Change between the Initial and Later Tests on the 1916 Revision

	1 and 2 (N = 32)	1 and 2 (N = 22)	1 and 3 (N = 22)	2 and 3 (N = 22)
a. Average percentages of change	6.88	7.8	8.9	4.6
b. Range of percentages of change	0-25.6	0-25.6	0-25.3	0-10.44
c. Range of gains based on the percentage of change	1.2-25.6	1.2-25.6	2.3-25.3	1.0-10.44
d. Range of losses based on percentage of change	1.0-16.0	1.0-16.0	1.1-12.0	2.0-8.5

exceeds that reported by Terman, although that for the second and third test on the smaller group is in approximate agreement.

In some instances our range for the percentage of change in *IQ* from one testing to another greatly exceeds that reported by Terman. Seven of the present subjects show percentages of change at

least double or more than the 8 per cent reported by him. Approximately 40 per cent of this group show percentages of change greater than the eight per cent.

The gains or losses computed on the basis of percentage of change do not appear to be limited to those above or below the average. One child of superior ability on the initial test gained 19 *IQ* points (16.5%) whereas another whose initial *IQ* was almost the same declined 14 points (11.9%). Two cases with initial *IQ*'s of 100 dropped to 84 on the second, and one with an *IQ* of 101 increased to 120. Another child at the borderline level increased her second *IQ* over the first by 16.2 per cent. However, the greatest number of those who show large changes in per cent, gains or losses, had initial *IQ*'s in the average or above average levels. This may be a selective factor since an attempt is made to choose children for placement who are at least fairly close to the average range.

Computation of the regression equation and the predicted scores show differences from the actual *IQ* on the second test based on ratios $Diff./PE_{(est.)}$ ranging from approximately 0 to 4.08. If a ratio of approximately 2.5 $PE_{(est.)}$ is used as a criterion suggestive of a change that is significant (this would include all cases where the *IQ* change is 15 or more points), five subjects gain and two decline.

With reference to those tested on Form *L* (Table 7), changes in *IQ* range from 0 to 5.92 times the $PE_{(IQ)}$ reported by Terman and Merrill (4) for the various *IQ* levels. Actual differences in *IQ*

TABLE 7

RATIO OF *IQ* CHANGES ON 53 SUBJECTS RETESTED ON THE SAME FORMS (*L* AND *M*) AS BASED ON $PE_{(IQ)}$ QUOTED BY TERMAN AND MERRILL FOR EACH *IQ* LEVEL

Ratio	Gain	Loss	No change	Total
0- .49	3	6	5	14
.50- .99	0	3		3
1.00-1.49	4	0		4
1.50-1.99	7	1		8
2.00-2.49	2	2		4
2.50-2.99	10	0		10
3.00-3.49	1	2		3
3.50-3.99	1	0		1
4.00-	6	0		6
	34	14	5	53

between the first and second test range from 0 to 18 points with an average change of 6.2. Changes on Form *M* range from 0 to 12.1 times the $PE_{(IQ)}$. Obtained changes range from 1 to 20 *IQ* points with an average of 6.0.

According to Terman and Merrill (4) the probabilities that an obtained score does not differ from the true score by more than 3 $PE_{(IQ)}$ are 22 to one. On this basis, 10 cases in the present data show changes where the probabilities of a difference run high. Of those gaining, six were initially in the average range and one was above and the other below. The two cases declining were within the average range, while 18.9 per cent of the total group tested on Forms *L* or *M*, then, show changes which appear to be significant.

The regression equation was computed for each form on the basis of the presented data. Ratios computed on the basis of $Diff./PE_{(est.)}$ range from .38 to 3.72 on Form *L* and from approximately 0 to 3.95 on *M*.

On the Form *L* initial test the deviations from the smoothed average *IQ* for each chronological age level computed by Terman and Merrill (4) range from —28.8 to +3.5 on the first testing and from —30.6 to +13.2 *IQ* points on the second. Changes with reference to approximation of the average *IQ* for each age level range from losses of 7.9 to gains of 16.7 *IQ* points. The range on Form *M* for the first test is from —34.9 to +34.1 and on the second from —25.4 to +38.7. Changes range from a loss of 4.8 to an increase of 19.2 *IQ* points. Of 11 changes in subjects above the average for their respective chronological level on the first test, nine show further increases and two decline somewhat.

Correlations between age at placement and *IQ* change, time interval between testing and *IQ* change, and between length of foster care and *IQ* change are all positive but too low for predictive purposes. Correlations between initial *IQ* and gain as computed by the formula given by McNemar (2) are +.346 for Form *L*, +.282 for *M* and +.155 for the 1916 Revision. These, however, may again reflect the selection of cases for foster placement since the original range is not great.

C. Analysis of Factors Contributing to the Changes Found in Individual Cases

Data from the case workers concerning the foster parents and homes are largely descriptive in nature. As previously reported (1),

the socio-economic status of the foster homes, however, tends to be above the average of the original homes. From data in the social histories the foster homes appear to be characterized by well planned programs of activities, relative freedom from emotional tensions, and with evidences of sincere interest in the children placed in their care. In contrast, the natural parental homes are usually tense, lack many constructive activities and interests for the children and tend to function on haphazard schedules. Many of the natural parents are alcoholic, sexually promiscuous, indifferent to their responsibilities, and a large number of the family backgrounds show incidence of insanity, chronic dependency, and delinquency.

In analyzing the underlying factors which appear to contribute to marked increases in *IQ*'s on these cases the following appeared to be of significance.

1. Freedom from previous home situations characterized by strong emotional tensions and a feeling of acceptance in the foster home situation.
2. Release from emotional tensions created by feelings of social rejection.
3. Strong needs for recognition, attention, and approval by others which now function in an area where intellectual pursuits are favorably recognized.
4. Availability of opportunities to achieve ambitions which were not possible in the natural homes either because of economic limitations or failure of the natural parents to encourage the interests.
5. Needs for a feeling of personal security and acceptance which had not been met by the natural parents but which the foster parents were capable of providing.

For those where there was a large decline in *IQ* the following seemed to be of significance:

1. An inability to accept the foster home placement and a "rebellion" against what were felt to be its unfair restraints.
2. A feeling of rejection and unfair criticism or discipline by the foster parents.
3. A desire to return to a previously unhappy home situation which was recognized as such but which the child continued to regard as more satisfying than the present home.

In all of these cases there was no single factor which one could say was *the* causal one. Rather, combinations of factors involving the child's previous feelings and those now operative in relation-

ship to the total patterns of the foster home situation appeared to be evident in each case where changes were found. In some instances, particularly the increases, the changes appear to have involved strong compensatory mechanisms which were substitutes for other symptoms being manifested at the time of the initial testing and interview. Changes of this type, then, would not necessarily indicate ideal re-adjustments in the personality structure but, instead, that the symptoms now function in a more socially acceptable manner.

The following brief case histories afford a picture of the changes found in some of the individual cases:

Feeling of rejection by the foster parents

R. S. on initial testing contained an *IQ* of 111 (1916 Revision). When re-examined two years later it was 97 and after the third year it was 110. (Because he exceeded the test ceiling, Form *M* was used about a week later when an *IQ* of 123 resulted.) At the time of the second test it was obvious that *R.* was unhappy although there was no direct evidence that the *IQ* itself was unreliable, and had there not been the previous test results it would not have been questioned. *R.* at this time described the foster home as a "free one, costing nothing" but was evasive as to actual relationships and feelings. He finally requested that he be moved to another home. Shortly after this was accomplished he was reexamined and described the previous home as "one where they had no confidence in him, all three women told him what to do and sometimes contradicted each other's directions; one of them often remarked that she would lose her mind if around him much; where perfection was wanted with no toleration of boyish interests." He also indicated that he was extremely anxious to complete High School and to graduate from college. Since the foster parents had frequently remarked about having helped another boy through college, *R.* thought it best to tolerate the conditions in the hope that he too might be so assisted.

Acceptance by foster parents

J. R. on the first testing (*M*) achieved an *IQ* of 92 and on the second one of 105. His original home was one where he was largely left to his own devices, roaming the streets and frequently in difficulties. The parents were divorced and when *J.* lived with his mother he was usually without supervision since the mother slept during the day and worked nights. The

father worked as a hospital attendant and was usually "too tired" to provide much companionship when *J.* stayed with him. In his present home the foster father is described as a vigorous, interested, well balanced person who enjoys spending considerable time with *J.* Despite being a busy farmer he is always willing to drop his work in order to discuss the boy and ways of helping him meet his various personality needs. During the second interview *J.* voluntarily remarked that his "present home is happier than when in his own." He would like to hear from his parents occasionally but felt that "the freedom from quarreling and fighting" had been good for him.

Inability to accept foster care

When first seen *C. L.*'s *IQ* (1916 Revision) was 117. The clinician then noted that the performance was almost entirely superior. On the second testing the *IQ* was 111 and on the third, 103. During both the second and third interviews *C.* spoke bitterly of his foster home placement, criticizing the foster parents, school teacher, and his social associates. He continued to cling to the hope of returning to his own home despite the fact that it continues to be unwholesome from the social viewpoint. Subtests on which he had previously been successful were failed on the second and third testing. Efforts at motivation within the limits of the test standardization were not well responded to.

Insecurity in the present foster home and a desire to return to a previously unpleasant situation

F. B. when first tested had an *IQ* of 107 (*M*) and on the second one of 92. Prior to coming under one of the agencies he had gone to live with a maiden lady who was interested in taking his sister into her home. She reluctantly agreed to "keep him" when he strenuously opposed separation from the sister. When first seen he presented problems of stealing, urinating out the windows and enuresis which had been met by threats of putting him in an institution, and physical punishment. During the second interview he frequently expressed a desire to return to his former home although he knew this was not possible.

D. Summary

One hundred and twenty subjects have received two or more psychometrics the results of which, in some instances, are suggestive of changes greater than would usually be expected. These differ-

ences between initial and retest *IQ*'s are usually found in subjects where the underlying emotional adjustments have also changed in either a favorable or increased in an unfavorable direction. The extent and direction of the changes found occur at each *IQ* level and do not appear to be a function of the initial *IQ*; those above or below the average level may either increase or decrease. Both the averages and the standard deviations of the distribution tend to increase in the second or later tests, although the differences are not always statistically reliable.

E. Conclusions

From these data it would appear that the foster home experiences and the individual's reaction thereto may be accompanied by changes from the initial *IQ*. Such changes can sometimes serve as *one* of the indices to the subject's adjustment in the new situation, suggesting whether or not satisfactory progress is being made.

References

1. Layman, J. W. A clinical study of children under foster care. *J. of Psychol.*, 1940, **10**, 107-120.
2. McNemar, Q. A critical examination of the University of Iowa Studies of environmental influences upon the *IQ*. *Psychol. Bull.*, 1940, **37**, 63-92.
3. Terman, L. M. The Measurement of Intelligence. New York: Houghton Mifflin, 1916. Pp. 362.
4. Terman, L. M., & Merrill, M. A. Measuring Intelligence. New York: Houghton Mifflin, 1937. Pp. 460.

Psychological Clinic
Indiana University
Bloomington, Indiana

The Journal of Genetic Psychology, 1942, 60, 71-84.

TENSION AND RESPIRATORY PATTERN IN YOUNG CHILDREN*

Clinic of Child Development, Yale University

MARY FRANCES WILLIAMS

A. THE PROBLEM

The present status of the problem of the rôle of tension can be summarized as follows. General over-reactivity in children and adults has been found by Rathbone (7), Lee (6), Wenger (8), and others to be characterized by chronic muscular hypertension. That this hypertension may be due to antagonistic "sets" operative in a conflict situation, is suggested by a conflict theory proposed by Castner (2). If this is true, tension would be expected to vary with the strength of opposing "sets," even in a non-emotional situation such as during performance of fairly simple tasks. Studies of changes in the respiratory pattern during mental activity show a tendency for a decrement in breathing amplitude to occur with increasing difficulty of task; this phenomenon may be directly related to the tension problem.

B. THE PRESENT EXPERIMENT

The aims of the present experiment are twofold: (*a*) to determine differences in tension, using various measures of the respiratory pattern as indices of tension, between easy and difficult activities, and between the beginning and end of both activity and inactivity periods; and (*b*) to determine the consistency of individual patterns of breathing.

This experiment is intended to test the adequacy of measures of breathing as indicative of what is known as low-level tension; although not subject to direct measurement, this has been in research frequently assumed to be correlated with muscle tension and so measured. By low-level tension is meant tension constituting a "set" for a given activity rather than the high-level extremes of tension characteristic of emotional states. When an individual is "set" for an

*Accepted for publication by Arnold Gesell of the Editorial Board, and received in the Editorial Office on September 9, 1940.

activity, a certain amount of mainly covert muscular activity is present, which is observed to increase with time if the stimulus for the overt act is not received immediately, or if the situation is such that, even when the stimulus is received, immediate response for any reason fails to proceed smoothly, especially if the act is a difficult task. With increase in local tension, there is manifest a concomitant increase in general tension, which tends to be maintained.

C. Subjects

Twenty-five nursery-school children of both sexes, ranging in chronological age from 36 months to 60 months inclusive, were used as subjects.

D. Apparatus

Since the standard pneumograph is unsuitable for work with young children, a less sensitive, but lighter and inconspicuous, rubber system was devised, as suggested by Professor Halverson. Two small, partially inflated balloons, inserted in a sleeveless cloth jacket suitable to be worn over the child's indoor clothing, were connected each to a rubber tube leading to a Marey tambour. The semi-inflated balloons were inserted each in a round sack fastened to the inner side of the jacket slightly below the thoracic and abdominal levels respectively of the tallest children. Lengths of belting, sewed to the jacket between jacket and sack at each level, were fastened by buckles behind the child. A wax-paper polygraph was used for recording.

In determining the validity of the experiment, it is necessary to determine the fidelity with which the record here obtained serves as an index of the amount of muscular activity in the respiratory regions. Assuming that normal respiration involves a reciprocal, rhythmical innervation and inhibition of antagonistic muscles, a decrease in respiration amplitude must be correlated with either a decrease in innervation of normally innervated muscles or an increase in the activity of those normally relaxed. It is assumed that an increase in innervation of normally relaxed musculature is indicated by a decrease in amplitude of the respiratory curve. The tube and dam system was calibrated by removing the balloon and inserting a syringe in the system at the point of connection between balloon and tube. The direct and near-linear relationship found between the amount of air in cubic centimeters and excursion of the stylus indi-

cates that such sources of error as stretching of the tubing, decrease in elasticity of the membrane with increased pressure (the Dodge type of dam was used, which was especially designed to eliminate this source of error), and arc distortion of the stylus would have a negligible effect in reducing the accuracy of the measure.

The length of record desired at one sitting obviated the possibility of using the more sensitive smoked paper on the ordinary kymograph drum. A signal-marker, which was in circuit with a 3½-volt battery and a switch near the examiner's right hand, was used to mark the beginning and end of each task or period of activity. A second electric marker, also in circuit with the battery and with an electric timer operating at one revolution per second, was used to mark one-second intervals on a time-line. The polygraph was operated at the rate of ½ inch per second; slight variation of the polygraph rate necessitated the use of a constant timer.

The clicking sound of the timer, which was in operation constantly from the time that the subject entered the room, served as a background noise which partially masked the sounds of the motor and switches. No persistent difficulties in adjustment to the situation were encountered.

A low chair and table were so placed outside the screen which concealed the polygraph and other apparatus that the child was surrounded on four sides by the screen, the back wall, the examiner's chair, and the table.

E. Materials

The following materials were used:

1. A story book (1), with illustrations in black and white, containing simple stories suitable for 3- and 4-year-olds.
2. A simple diagram or path, 16 inches in length, graduated in width from ½ inch to ⅛ inch, colored brown on a green paper background.
3. Twenty-six uncovered paper boxes, 2 inches square, fastened onto cardboard stands in groups of 8, 4, and 1; all boxes in each group were of a different hue or so markedly different in saturation as to cause the subjects no difficulty in matching paper discs to them.
4. Round paper discs 1½ inches in diameter to be sorted into boxes according to color, 8 discs for each group of boxes.
5. Two small (2-inch) dolls.
6. (*a*) A wooden "bridge" 9 inches in length, made with two

wires pulled taut in a wooden frame (10" x 4"), the distance between the wires graduated from ¼ inch to ¾ inch, with six ⅝-inch rectangles of paper glued at equal intervals along the "bridge" forming steps of graduated width; (*b*) a toy man (1½ inches high) fastened on a flat base (1 inch square).

7. Colored crayons and white drawing paper (5½" x 8½").
8. Stop watch.
9. A round desk bell.
10. A "theatre" consisting of a roll of white paper passing between two round slotted containers, with pictures glued onto the paper at intervals of 24 inches.
11. Ten one-inch wooden cubes, painted orange.

F. Procedure

The procedure of the examination of each child was as follows. The child was brought into the room by the examiner, with these instructions: "*Here is where we play. That is your little chair and this is my chair, and this is the little apron that you put on. We'll put it on before you sit down.*" If the child asked any questions about the jacket, he was told briefly that all the children wear it when they play with the things in that room. Every effort was made to make the initial adjustment as natural as possible; the jacket was arranged on the chair before the child entered the room in such a way that the balloons and tubing were concealed. In adjusting the jacket, the upper belt was fastened as high as possible under the arms, the other at waist level. As adjustment had to be made with care, a small picture book was handed to the child to look at during this time. No children showed persistent interest in the jacket.

Two forms of the examination were given to each child, with at least four days intervening between successive examinations of the same child. One-half of the subjects were given Form *X* followed by Form *Y*; the other half of the group were given Form *Y* followed by Form *X*. The two forms were identical in content; the order of tests was reversed in the second form, thus partially controlling the factor of fatigue. The two short stories of two pages each were first read to the child, the reading of which took approximately three minutes, previous trials having shown this length of time to be sufficient in all cases to give relaxation as indicated by

regular breathing, regardless of the degree of activity on the playground in which the subject was engaged just before the examination.

The tasks used fall into three groups, those in which (*a*) tasks graded in difficulty from easy to difficult, and vice-versa, were included (Tracing; Sorting the Bridge; Drawing; Tower); (*b*) tasks in which a given situation was prolonged presumably longer than desired by the child (Tapping); and (*c*) inactivity periods in which the child awaited a signal (Guessing; Bell-ringing). The number of tasks included under each of the three headings is consistent with the tediousness for the child of that type of task.

Tower-building was selected as a good example of a task in which, with the increasingly fine coördination required, tension increases rapidly. In order to secure a comparable task in which the chief requirement should be fine coördination in balancing blocks, but in which tension would be expected to decrease from a high point at the beginning of the task to a low point at the end, the "bridge" was devised. In this way any differences found between periods of high and low tension would necessarily be attributable to a factor other than fatigue. In all other tasks, reversal of order was possible, so that a given task could be presented in easy-to-difficult or difficult-to-easy order.

The procedure of the examination for each task is described briefly below, in the order in which tasks were presented in Examination Form *X*.

I. *Tracing.* Following demonstration by the examiner, the subject was instructed to trace the path beginning at the wide end.

II. *Sorting.* Groups of 8, 4, and 1 boxes were presented successively; the subject was instructed to sort the colored discs into appropriate colored boxes.

III. *Guessing.* The child was instructed to place his hands on the table behind a piece of cardboard, and to close his eyes until he should feel an object (doll) placed in his hands (20" period); then to guess what was in his hand.

IV. *The Bridge.* The toy man was to be balanced on each successive step of the bridge, beginning at the narrow end, following demonstration.

V. *Drawing.* The subject was instructed to copy (*a*) a square, (*b*) a square crossed by diagonals, (*c*) a square crossed by diagonals, vertical and horizontal mid-lines.

VI. *Tapping.* Instructions were to begin tapping with the

crayon on paper; instructions were repeated after each of the first three spontaneous refusals or pauses in tapping. Thus breathing during a total of four periods of tapping was recorded.

VII and VIII. Same as Tasks I and II respectively, each task presented in reverse order of difficulty.

IX. *Bell-ringing.* The bell and roll were presented on the table. The child was instructed to ring the bell as soon as he saw a picture appear on the roll which the examiner operated by hand. Intervals were of 10".

X. *Tower-building.* The subject was instructed to build a single tower of 10 blocks.

XI and XII. Same as Tasks V and VI respectively, each task presented in reverse order of difficulty.

The above list of 12 tasks constituted Session *X*; the same tasks presented in reverse order, following reading of the same stories formed Session *Y*. Each session took approximately 30 minutes.

Instructions for each task were standardized.

It may be objected that even in the case of young children voluntary inhibition of breathing may have occurred in order to eliminate such secondary arm movements as are likely to occur, in the same way that the adult tends to hold his breath in performing a task requiring fine motor control. It was found, however, that difficulty of task was similarly correlated with amplitude of breathing even when the task involved a choice reaction as in sorting, rather than fine coördination.

G. Results

In determining the measures which would best differentiate between tension and non-tension periods, four objective measures of the respiration curves were initially taken; those measures were used which previous experimenters had found most adequate. The measures selected were: (*a*) amplitude, the summated values in oscillometer (5) units of vertical distances covered by inspiration and expiration phases during a given number of (five) full respiratory cycles; (*b*) regularity, summated values of vertical distances between successive crests of the respiration curve, over the given number of full cycles; (*c*) rate, the number of full cycles and fractions thereof above ½ per minute, taken from 10-second samples; (*d*) level, the median of vertical distances from an arbitrary baseline to the crests of the curve, over the given number of full cycles.

For discussion and illustration of common types of breathing, see Halverson (4).

The amplitude measure was found to be superior to the other three measures in differentiating between tension and non-tension periods. No significant differences in regularity or level were found (Table 1).

TABLE 1

BREATHING CHANGES WITH INCREASING DIFFICULTY OF TASK (TOWER-BUILDING)

Measure	Initial record	Final record	D	D/σ_D
Amplitude (oscillometer units)	32.5	23.5	—9.0	5.0
Rate (per minute)	43.8	51.0	+7.2	3.2
Regularity (oscillometer units)	4.3	3.4	—0.9	1.0
Level	5.7	5.7	0.0	0.0

Since breathing rate and amplitude were found to be the most adequate differential measures in a comparison of breathing patterns under varying conditions, further analysis of the records was made in terms of these two measures.

Breathing rate was measured during performance of all five tasks of graded difficulty and all three homogeneous activities. Results in terms of mean breathing rate per minute during performance of easy and difficult parts of the graded tasks and during the first and last 10-second periods of the homogeneous activities, differences, and critical ratios are given in Table 2.

In four out of five graded tasks, there is no marked change in breathing rate with either an increase or a decrease in the difficulty of the task; differences are not consistent or significant. During one graded task (the Bridge) there is marked decrease in breathing rate with decreasing difficulty of task during Session *Y*; with no significant change in rate during Session *X*. Since this task is the 4th activity in Session *X* and the 9th activity in Session *Y*, the results suggest that the response to this task is particularly affected by fatigue. Since, however, breathing rate is not in the case of the other tasks generally differentially affected by the time of presentation of the task within the session, the importance of this factor is questionable. Breathing rate changed significantly during only one session of a homogeneous activity (Bell-ringing *Y*) ; this change was in the opposite direction from the rate change during the Bridge

TABLE 2

MEAN RESPIRATORY RATE (PER MINUTE) DURING PERFORMANCE OF TASKS

	Task	Mean rate Easy	Difficult	Difference	D/σ_D
Sorting	X I	45.6	44.2	—1.4	0.7
	X II	50.0	43.4	—6.6	2.6
	Y I	47.6	50.4	+2.8	1.0
	Y II	50.9	51.0	+0.1	0.1
Drawing	X I	48.5	45.1	—3.4	1.6
	X II	46.4	48.7	+2.3	0.9
	Y I	48.4	44.4	—4.0	1.5
	Y II	45.9	52.0	+6.1	2.5
Tower	X	45.6	46.8	+1.2	0.9
	Y	43.8	51.0	+7.2	1.2
Bridge	X	42.5	43.1	+0.6	0.3
	Y	46.7	36.8	—9.9	5.9*
Tracing	X I	42.3	40.6	—1.7	1.0
	X II	42.4	43.8	+1.4	0.6
	Y I	45.8	43.7	—2.1	1.0
	Y II	46.2	45.1	—1.1	0.6
		Initial	Final		
Guessing	X	32.6	36.4	+3.8	2.8
	Y	42.8	41.9	—0.9	0.4
Bell	X	43.7	46.3	+2.6	0.8
	Y	42.3	50.6	+8.3	3.9*
Tapping	X I	45.6	46.8	+1.2	0.9
	X II	49.5	48.6	—0.9	0.4
	Y I	48.8	51.0	+2.2	0.8
	Y II	53.6	56.6	+3.0	1.6

*Statistically significant difference, i.e., critical ratio above 3.0.

task. The fact that only two out of 24 activity periods showed change in breathing rate, and that the two changes were in opposite directions suggests that breathing rate is in general little affected by a change in the degree of conflict involved in the nature of the activity.

The initial and final 10-second periods of all tasks were compared in terms of amplitude. During all but one of the tasks a decrement in breathing amplitude was found to accompany increased difficulty. During two out of four performances of the Tracing task, slight and insignificant increments were found. During the final 10-second periods of all homogeneous activities, in which opposing sets were expected to be operative at the end, breathing was

TABLE 3

RESPIRATION AMPLITUDE (ABDOMINAL) DURING PERFORMANCE OF TASKS

	Task	Mean amplitude Easy	Difficult	Difference	D/σ_D
Sorting	*X* I	34.3	25.6	— 8.7	3.9*
	X II	27.9	17.8	—10.1	3.3*
	Y I	28.6	22.4	— 6.2	3.3*
	Y II	31.3	24.1	— 7.2	3.2*
Drawing	*X* I	25.8	24.5	— 1.3	0.5
	X II	39.2	36.6	— 2.6	0.5
	Y I	36.2	32.7	— 3.5	1.1
	Y II	32.7	30.9	— 1.8	0.6
Tower	*X*	32.5	23.5	— 9.0	5.0
	Y	30.6	25.5	— 5.1	3.6*
Bridge	*X*	24.2	18.8	— 5.4	5.8*
	Y	29.6	22.2	— 7.4	3.2*
Tracing	*X* I	23.2	24.1	+ 0.9	0.4
	X II	27.5	33.5	+ 6.0	1.7
	Y I	32.7	29.9	— 2.8	0.8
	Y II	27.0	24.2	— 2.8	1.4
Guessing	*X*	38.8	29.0	— 9.8	5.1*
	Y	34.0	24.3	— 9.7	4.2*
Bell	*X*	27.6	22.6	— 5.0	1.1
	Y	30.4	27.0	— 3.4	1.0
Tapping	*X* I	30.0	21.0	— 9.0	4.1*
	X II	31.9	19.8	—12.1	3.5*
	Y I	34.2	24.8	— 9.4	4.5*
	Y II	32.9	25.2	— 7.7	5.1*

*Statistically significant difference, i.e., critical ratio above 3.0.

likewise depressed. Twenty-three subjects were used in determining these differences.

The fact that depressed breathing was characteristic of difficulty in the part of the task being performed, whether the task was of progressively increasing or decreasing difficulty shows that the difference in amplitude is due to another factor than either fatigue at the end of a task or initial excitement followed by calming down toward the end.

A reliability measure was taken by determining the rank-order correlation in terms of amplitude changes between the first and last five respiration cycles, from records of respiration during the tower-building task of Sessions *X* and *Y*. A rank-order correlation of .66±.08 (equivalent to a Pearson *r* of .68) was obtained.

The fact that correlation is low and insignificant (rank-order correlation: r, .09; PE, .15) between decrease in amplitude of respiration during tower-building, and chronological age, points to the conclusion that the obtained individual differences are to be attributed to factors other than that of maturity or motor coördination, which is known to be highly correlated with chronological age below five years.

The importance of the fatigue factor in determining breathing pattern changes was measured by a comparison of breathing amplitude changes during activities in the first and second halves of a session, during both sessions. Results are shown in Table 4.

TABLE 4

CHANGES IN RESPIRATION AS A FUNCTION OF FATIGUE

	Number of activity periods	Mean amp. change (Oscillometer units)	D	D/σ_D
First half session	12	—5.8	0.4	0.6
2nd half session	12	—5.3		

Over a period of approximately 30 minutes, the effect of fatigue upon changes in the breathing pattern in children of this age is negligible.

Objection might be raised that with virtual inhibition of respiratory activity in all children, a measure of the degree of change in respiratory amplitude occurring between a tension and a non-tension task, would resolve itself into a measure of the initial amplitude of breathing during the non-tension task, this amount being largely dependent upon such physical factors as body weight. A correlation by the rank-order method of —.18 (PE .14) between amplitude of respiration during the initial 10 seconds of tower-building and the decrease in amplitude with increased difficulty indicates that this factor is of little importance.

Increments in costal breathing amplitude were rarely so great as to compensate for decrements in abdominal breathing during performance of difficult tasks, or at the end of such activities as Guessing and Tapping. It can be assumed that the decrement in amplitude of abdominal breathing is not attributable to a change in the region of maximal activity from predominantly abdominal to predominantly costal breathing.

Mean amplitude and rate of breathing were determined for three

categories of behavior: (*a*) passive participation, listening to stories (mean of breathing measures from *X* and *Y* sessions); (*b*) active participation, performance of an easy task (mean of measures on easy parts of all graded tasks in both sessions); (*c*) active participation, performance of a difficult task (mean of measures on difficult parts of all graded tasks in both sessions). Results were tabulated as shown in Table 5.

TABLE 5

RESPIRATORY AMPLITUDE AND RATE FOR THREE CATEGORIES OF BEHAVIOR

Behavior category	Mean amp. (Oscillometer units)	Mean rate (Per min.)
Passive participation	28.6	41.9
Active participation (easy)	31.0	46.2
Active participation (difficult)	25.4	45.0

H. DISCUSSION

It is postulated that tension may alternatively be attributable to the presence of a conflict between a positive response to the situation and withdrawal from it, or to a conflict between two or more possible incompatible responses. Data from studies of the knee-jerk by Courts (3) and others indicate that activity in one set of muscles tends to facilitate muscular activity in other parts of the body. Courts demonstrated that the amplitude of the knee-jerk bears a linear relationship to the amount of pressure exerted on a hand-dynamometer. An increase in local muscle tension due to conflicting sets may be expected to spread to other musculature, including that involved in respiration.

An inverse relationship between the amplitude of respiration during performance of a task and the degree of conflict involved in performance of the task indicates a change in the normal pattern of innervation of the costal and abdominal musculature. The decrement in amplitude of breathing which occurs with increasing difficulty in the task being performed could be due to (*a*) relaxation of normally innervated musculature, or (*b*) innervation of normally relaxed musculature acting antagonistically upon that normally active. The majority of the experimental findings on the subject indicate that an increase in mental activity is accompanied by an increase in muscular activity, to such an extent as to indicate that a general state of muscle tension is probably present. It is justifiable,

therefore, to assume that a decrement in breathing amplitude is not attributable to decreased tension in the costal and abdominal musculature but rather to an increase in activity in these regions.

Since the net effect of general tension in the abdominal musculature amounts to a brief partial paralysis of these muscles, it may be objected that a decrease rather than the obtained increase in simultaneous dynamometer pressure with increased difficulty of task would be expected in studies of this type. If, however, the general level of dynamometer pressure is increased, this would constitute a change comparable to a change in the level rather than in the amplitude of respiration.

If tension is defined as a "feeling of strain or suspense" concomitant with the establishment of incompatible motor sets, then measurement of the strength of the opposing sets will be expected to constitute an index of the tension present in the organism, since the tension itself is known only to introspection and is therefore not directly measurable. If tension is defined operationally as "a state produced in an organism when a normally adequate stimulus is prevented from eliciting an immediate goal response" (2), a state involving essentially muscle tension, then a muscle tension measure can be taken as a relatively direct measure of tension.

In a tension theory logically developed from the latter definition it is assumed that when there is a barrier to any response for which a set has been established, general tension occurs, either primarily due to the situation or secondarily due to a spread of tension from the local region in which the set was established. As the stimulus and the barrier continue to operate antagonistically, the tension increases in degree. Further development of the theory covers the effect of local tension reduction upon general tension.

The present experiment is concerned with the increase in general tension concomitant with antagonistic operation of a stimulus and a barrier in various kinds of conflict situations. It is found that with an increase in the strength of a barrier due to (*a*) increase in precision of response required (e.g., Tower-building); (*b*) increase in the strength of a set antagonistic to a given activity in progress (e.g., Tapping); or (*c*) increase in the strength of a set for activity during an inactive period (e.g., Guessing), there is a concomitant increase in general tension as determined by an index of respiratory activity.

Further study of muscle tension in varying degrees of both mild conflict and emotional situations promises to throw light on the

mechanisms of response to these situations, as well as to make clear the actual rôle of tension, whether as an essential factor in the response, or as a by-product significant only as an index of more intrinsically significant events. The effect of a state of tension upon the nature of substitute responses, in situations where a prepared or set response is barred, has been barely touched upon in previous experiments. Study of the actual events occurring within tension systems, and of interaction of systems, can be expected to increase the predictability of human behavior in conflict situations. The relative value of various measures of tension, including the measure of respiratory activity used in this experiment, is also still to be determined.

I. Summary

Records were made of costal and abdominal breathing movements of 25 nursery school children (*CA* 36-60 months) during periods of inactivity and varying degrees of quiet activity. The following results were obtained:

1. Measures of general level, of regularity, and of rate of respiration failed to differentiate between periods of performance of easy and difficult tasks, or between degrees of conflict present in a homogeneous task.
2. Breathing amplitude decreased progressively during periods of homogeneous (*a*) activity, and (*b*) inactivity preeeding a signal for activity.
3. Individuals tended to be consistent in the amount of change in breathing pattern shown with changes in activity.
4. No relationship was found between changes in breathing amplitude and chronological age, or amplitude of relaxed breathing. Fatigue was found to have a negligible effect on these changes within the 30-minute experimental period.
5. During performance of tasks, a slight but not fully compensatory increment in costal breathing occurred concomitantly with the decrement in abdominal breathing amplitude as difficulty of task increased.
6. Mean amplitude and rate of breathing were determined for three categories of behavior: (*a*) passive participation in an activity; (*b*) active participation in an easy task; (*c*) active participation in a difficult task.
7. Results were discussed in connection with a theory of tension.

References

1. Baruch, D. The two Bobbies. New York: John Day, 1930. Pp. 124.
2. Castner, B. M. Muscle tension and variability of response pattern. (Not yet published.)
3. Courts, F. A. The knee-jerk as a measure of muscular tension. *J. Exper. Psychol.*, 1939, **24**, 520-529.
4. Halverson, H. M. Variations in pulse and respiration during different phases of infant behavior. *J. Genet. Psychol.* (In press.)
5. Hull, C. L. An instrument for summating the oscillations of a line. *J. Exper. Psychol.*, 1929, **12**, 359-361.
6. Lee, M. A. The relationship of the knee-jerk and standing steadiness to nervous instability. *J. Abn. & Soc. Psychol.*, 1931, **26**, 212-228.
7. Rathbone, J. L. Residual Neuromuscular Hypertension: Implications for Education. New York: Columbia Univ. Press, 1936. Pp. 220.
8. Wenger, M. A. Some relationships between muscular processes and personality and their factorial analysis. *Child Devel.*, 1938, **9**, 261-276.

Clinic of Child Development
14 Davenport Avenue
New Haven, Connecticut

The Journal of Genetic Psychology, 1942, **60**, 85-98.

A COMPARISON OF THE KUHLMANN-ANDERSON INTELLIGENCE TEST WITH THE REVISED STANFORD-BINET, FORM *L**

Child Guidance Clinic of the Sunbeam Home Association, Oklahoma City, Oklahoma

THEODORE CARLTON[1]

In the Kuhlmann-Anderson instruction manual (12, p. 28), the authors suggest that their tests may be used for an individual examination. Many psychologists have adopted this suggestion and are today employing this test in place of the Revised Stanford-Binet (16) or the Kuhlmann *Tests of Mental Development* (11). Since we know very little about the relationship of the Kuhlmann-Anderson with the 1937 Stanford or the 1939 Kuhlmann, such a procedure is questionable.

When an examiner substitutes one test for another, he automatically makes two assumptions: first, that both tests will yield the same *IQ* or mental age scores over the entire range of measurement, or at least will not differ by the PE_{IQ} of the one chosen as the standard; and second, that both examinations will discriminate to the same extent between different *IQ* or mental age ranges. In the past, educators, publishing houses, and clinics have accepted these hypotheses without proof, believing that if all other conditions of a good examination were present, an *IQ* on one test was equal to the same *IQ* on any other; and that any well standardized intelligence test could be used to measure all degrees of mental deficiency or superiority.

In 1924, Miller (15) designed an experiment to question this general hypothesis. Within a short space of time, he gave nine group intelligence tests and the Stanford-Binet, 1916 Revision, to 57 high school freshmen. After analyzing the results, he concluded (p. 366) that "the mental age norms vary so much that it is impossible to interpret the *IQ*'s from all group tests according to the

*Accepted for publication by L. M. Terman of the Editorial Board, and received in the Editorial Office on September 27, 1940.

[1]The writer wishes to thank Dr. E. J. Engberg, Superintendent of the Minnesota School and Colony, for permitting and encouraging the research; and L. E. Carlton who administered all of the Kuhlmann-Anderson tests.

Stanford-Binet standards." Kefauver (8), after administering 12 different group mental tests to 100 high school pupils, corroborated Miller's conclusions. In addition, he demonstrated that the variation between the scores on different tests was greater at the extremes of the distribution.

Since 1929, Kuhlmann (10), Hilden and Skeels (7), Hartman (6), and Carlton (1, 2), have obtained similar results on different populations. Kuhlmann examined 1,400 Minnesota children in Grades I through XII with eight group mental tests. He found differences not only between the mean *IQ*'s at each grade level, but also in the power of the tests to discriminate between grade levels.

While most of the other writers had experimented with group intelligence tests, Hilden and Skeels compared the Kuhlmann-Anderson Test with the 1916 Stanford-Binet. After making a thorough analysis of the test results of over 700 subjects between the ages of 6 and 21, they were able to show that some children varied from one test to another by as high as 28 *IQ* points. On the basis of their results, they constructed regression equations.

Recently Hartman (6) compared the 1916 Stanford-Binet with the Bregman Revision of the Army Alpha. His subjects were 1,443 adult male prisoners who had scored between 0 and 24 points on the Alpha. After examiners had also tested these cases with the Binet, he found (p. 611) that the "Alpha test does not measure mental defectives distinctly enough for reliable diagnosis. . . ."

Within recent months, the writer has carried out two experiments to test the hypothesis that the Kuhlmann Tests (11) and the Revised Stanford-Binet (16) were comparable at various age levels. Employing a total of 186 mental defectives between the ages of 7-6 and 17-5, we found that significant differences between mean *IQ*'s occurred at various levels. We concluded that experimenters were unjustified in substituting one test for another and assuming that both would give the same results.

Procedure of This Experiment

This research was planned to test two hypotheses: First, that *IQ*'s on the Kuhlmann-Anderson Test are equivalent to *IQ*'s on the 1937 Revised Stanford; and second, that both tests will discriminate between different *IQ* ranges to the same extent.

In order to carry out this experiment, we first chose 112 children

at random from a previously selected group at the Minnesota School and Colony.[2] Then, employing the Kuhlmann-Anderson *Intelligence Tests* as an individual examination, one experienced examiner administered them to the group. In giving these tests, we followed carefully the instructions of the manual (12, p. 28). Thus, each subject took the 14 sub-tests that were best adjusted to his mental level; that is, if a subject scored a zero on a sub-test at the beginning or at the end of the test, these were eliminated and he was given easier sub-tests from the earlier levels. Since the Kuhlmann Tests and the Revised Stanford-Binet were also being administered during this time, we were careful, first, that the administration of the Kuhlmann-Anderson and the other tests were separated by a period of at least one month; and, second, that half of the group of 112 was tested with the Stanford[3] before the *K-A*. The difference in time between the two tests ranged from one to four and a half months with the mean at 1.8.

LIMITATIONS OF THE PROBLEM

Since our main purpose was to compare the relationship of the Revised Stanford-Binet with the Kuhlmann-Anderson, used as an individual examination, we had to set up one test as a standard for the other. It is conceivable that, on the basis of subjective choice, one could select either of these tests as a standard. In this case, however, we are comparing the *K-A*, which was standardized on groups of children, with the Stanford which was standardized on individual children. Since the *K-A* is primarily a group test, it must conform to the standards of the Stanford if it is to be used as an individual examination.

Another difficulty which faced us was the different method which

[2]In a previous experiment (2), we had selected all children between the ages of 7-6 and 15-11. We eliminated from this group all who had serious speech, physical, or sensory handicaps, and who could be classified as a "clinical type" (9, p. 202). We then administered the Revised Stanford-Binet, Form *L* to half (*A*) of the 154 remaining subjects and the Kuhlmann *Tests of Mental Development* (11) to the other half (*B*). Group *B* was then given the Stanford and *A* the Kuhlmann. Following this, we tested all who could read sentences with the Gray *Oral Reading Paragraphs Test* (5); and if a child's reading age was below 8-6 and was eighteen or more months behind his adjusted mental age (to obtain this value, the Revised Stanford-Binet was rescored so that the influence of the reading tests was removed), he was excluded from the study.

[3]The Stanford was scored to failure on two successive levels.

each author used to compute the *IQ*. Whereas Terman and Merrill (16, p. 30), between 13-0 and 16-0 "cumulatively drop one out of every three additional months of chronological age and all of it after sixteen" in computing the *IQ*. Kuhlmann and Anderson use the exact chronological age as a divisor up to 16-0.[4] Since we computed the *IQ*'s according to each author's recommendation, we had to divide the present sample into two parts: First, the *CA* group between 7-6 and 13-1; and, second, the group between 13-2 and 15-11.

Results

In Tables 1 and 2 we have listed the essential *IQ* and mental age data for each test and age group. After inspecting the mean *IQ*'s and *MA*'s, one will note that, in each case, the mean differences are in favor of the *K-A*. When we applied the Fisher "*t*" test for paired variates (4, p. 42), we found that the differences between the means were significant at the 1 per cent level for Group *X* in Tables 1 and 2. We may be practically certain that the true mean differences are, in these cases, greater than zero and lie between the fiducial limits (4, p. 39-40) listed in the tables.

Although the differences between the means for *Y* group, Tables 1 and 2, are insignificant, we cannot conclude that the *IQ*'s on the two tests are entirely equivalent. For this to be true, the two tests must satisfy three additional criteria. First, their *sigmas* and ranges must be identical, second, the *K-A* at any point in the range under consideration, must not differ from the Stanford by more than PE_{IQ} on the Stanford; and third, the correlations between the two tests must be close to +1.00 and significant at the 1 per cent level.

From Table 1 we note that the ranges and the *sigmas* in the *Y* group are not identical. In fact, the *sigma* for the Stanford *Y* group differs positively from the *K-A* and shows significance at the 1 per cent level.

By calculating the average differences between the *IQ*'s which subjects scored on Forms *L* and *M* of the Revised Stanford, Terman and Merrill have derived a PE_{IQ}. According to these authors (16, p. 46) the PE_{IQ} between the *IQ* range of 70 and 89 is ± 2.60

---

[4]To facilitate the computation of *IQ*'s, Terman and Merrill have provided an extensive table (p. 415 ff.). Table VI in Kuhlmann (11, p. 233 ff.) may be used for the computation of *K-A IQ*'s.

TABLE 1

INTELLIGENCE QUOTIENT DATA FOR 112 SUBJECTS WITH *CA*'S FROM 7-6 TO 15-11 AT THE MINNESOTA SCHOOL AND COLONY

Intelligence test	No. of cases	Chronological age range	Mean *CA* (mos.)	*IQ* range	Mean *IQ*	M_{K-A}-M_S	"t"$_D$*	Fiducial limits of means 1% Level	σ_{IQ}	"F"**
Revised Stanford, Form *L*	62	*X* 7-6 to 13-1	131.85	46 to 90	64.85	+3.36 (*K-A*)	5.17 $P<.01$	+1.90 to +4.82	9.72	1.20 $P>.05$
Kuhlmann-Anderson	62		132.53	51 to 91	68.21				8.88	
Revised Stanford, Form *L*	50	*Y* 13-2 to 15-11	175.52	45 to 100	62.72	+1.20 (*K-A*)	1.01 $P>.10$	−1.98 to +4.38	11.88	2.32 $P<.01$
Kuhlmann-Anderson	50		175.10	46 to 80	63.92				7.80	

*Fisher t test for difference between means, paired variates.
**F ratio of the variances, $\sigma^2_S/\sigma^2_{K-A}$. A test of the significance between two *sigmas*.

TABLE 2

MENTAL AGE DATA FOR 112 SUBJECTS WITH *CA*'S FROM 7-6 TO 15-11 AT THE MINNESOTA SCHOOL AND COLONY

Intelligence test	No. of cases	Chronological age range	Mental age range	Mean mental age	M_{K-A}-M_S	"t"$_D$*	Fiducial limits of means 1% Level	Pearson r of MA's	t_r**
Revised Stanford-Binet, *L*	62	*X* 7-6 to 13-1	4-11 to 9-8	84.68	+4.63 (*K-A*)	5.45 $P<.01$	+2.37 to +6.89	.847	12.65 $P<.01$
Kuhlmann-Anderson	62		4-11 to 10-0	89.31					
Revised Stanford-Binet, *L*	50	*Y* 13-2 to 15-11	6-2 to 14-7	106.00	+4.70 (*K-A*)	$P>.01$ $<.02$	−.54 to +9.94	.763	8.29 $P<.01$
Kuhlmann-Anderson	50		6-2 to 12-2	110.70					

*Fisher t test for difference between means, paired variates.

**Fisher t test of significance of the correlation coefficient.

and below 70 is ± 1.49. In order to be interchanged, IQ's on the $K\text{-}A$ would then have to satisfy the criteria in Table 3. An examination of Table 5 will show that these conditions are satisfied in only the X_1 group.

From a relative standpoint the correlations between mental ages on both tests should at least reach $+.91$ or lie between the range $+.91$ and $+.98$ (16, p. 7). That this is not the case is evident from Table 2.

We doubt, therefore, whether an examiner is justified in substituting the $K\text{-}A$ for the Stanford, believing that they will yield the same IQ.

We stated earlier that Hilden and Skeels (7) had compared the Kuhlmann-Anderson with the 1916 Stanford-Binet. Since they drew their subjects from institution populations and since the CA's of their subjects extended from 6 to 21, their investigation was somewhat similar to ours. It was dissimilar in that they employed the $K\text{-}A$ as a group test, used a wider range of mental ability, and the 1916 Revision of the Stanford. In spite of these differences, it is noteworthy that they concluded (p. 229) that the $K\text{-}A$ test tended "to yield slightly higher IQ's than the Stanford-Binet scale." This was shown "by the average difference of $+1.6$ points on the obtained IQ's and by the average difference of $+2.4$ and $+0.5$ on estimated values."

TABLE 3

PER CENT OF CASES ON *KA* THAT MAY DIFFER FROM THE STANFORD BY NOT MORE THAN THE LISTED LIMITS

IQ Ranges	Limits in terms of *IQ* Points			
	More than ± 5	0 to ± 5	0 to ± 4	0 to ± 3
70 to 89	19.6	80.4	70.1	56.2
46 to 69	2.4	97.6	92.9	82.4

DISCRIMINATING ABILITY OF EACH TEST

Psychometrists very often employ an intelligence test to aid in the diagnosis of mental deficiency. In doing this they are interested in using an instrument that will definitely discriminate between small IQ ranges at the moron and imbecile levels. To what extent, then, will the Revised Stanford and the Kuhlmann-Anderson make this discrimination? Table 1 gives us only a small clue to a possible

answer. One will note that the difference between the *sigmas* in the Y group is in favor of the Stanford. When we determine the "F" ratio of the variances, that is, $\sigma^2_S / \sigma^2_{K\text{-}A}$ (13, p. 60), we find that it is significant at the 1 per cent level.

In a general way, this indicates that the two tests do differ in their powers of discrimination. In order to obtain a more accurate measure of any differences which may exist, we divided each large group, X and Y, into two sub-groups, a and b. In Table 4, X_a refers to the 33 subjects between the *CA*'s of 7-6 and 13-1 who scored *IQ*'s between 46 and 64 on the Revised Stanford-Binet; and X_b refers to the 29 who scored *IQ*'s between 65 and 90; and so on.[5]

We then computed the differences between the Revised Stanford and the Kuhlmann-Anderson *IQ*'s ($K\text{-}A_{IQ}$'s — S_{IQ}'s) for each individual. When we calculated these mean gains for all sub-groups, we found the following results (Table 4):

1. For the X_a and Y_a groups the mean gains in *IQ* points are significant at the 1 per cent level, will differ positively in favor of the *K-A*, and will lie between the limits $+2.83$ and $+7.59$ (X_a) and $+2.69$ to $+7.91$ (Y_a).

2. Using the formula for distinct samples (4, p. 40), it becomes clear that the X_a group will gain a significantly greater number of *IQ* points than the X_b; and the Y_a will gain a significantly greater number than the Y_b group. When we inspect the fiducial limits for these mean differences, we find that the Y limits are much higher and much wider than the X.

3. If the mean gain ($+5.21$) of the X_a sub-group is added to the mean *IQ* (57.36) of X_a, the resulting value, 62.57,[6] approaches the mean *IQ* (64.85) of the total Stanford X group in Table 1. The same total mean value may be computed for the X_b group ($73.17 + 1.45 = 74.62$), for the Y_a ($53.93 + 5.30 = 59.23$), and for the Y_b ($72.87 - 4.52 = 68.35$). While the Y_a and Y_b values approach the mean of the Stanford group in Table 1, the X_b group does this only to a very slight extent.

Before concluding, we constructed Table 5. Here we note, in greater detail, the extent to which subjects differed on the Kuhlmann-Anderson from the Revised Stanford. After inspecting this table and Table 4, it becomes clear that the two tests differ rather

[5]The mean *IQ* that each group, X and Y, made on the Stanford was chosen as the division point.

[6]This is the mean Kuhlmann-Anderson *IQ* of the X_a group.

TABLE 4

GAINS IN *IQ* POINTS MADE BY THE KUHLMANN-ANDERSON OVER THE STANFORD: ANALYSIS OF FOUR SUB-GROUPS

C.A Range	Sub-class	No. of cases	IQ's on Revised Stanford	IQ points gained by the K-A over the Stanford: Range	Mean*	t_M	Fiducial limits of means 1% Level	M_a-M_b	t_D**	Fiducial limits of differences 1% Level
7-6 to 13-1	X_a	33	46 to 64 M=57.36	— 4 to +14	+5.21	5.99 P<.01	+2.83 to 7.59	+3.76 (K-A)	3.03 P<.01	+ .46 to +7.06
	X_b	29	65 to 90 M=73.17	— 8 to + 9	+1.45	1.65 P>.10	— .81 to +3.71			
13-2 to 15-11	Y_a	27	45 to 62 M=53.93	— 4 to +21	+5.30	5.64 P<.01	+2.69 to 7.91	+9.82 (K-A)	4.89 P<.01	+4.43 to +15.21
	Y_b	23	63 to 100 M=72.87	—26 to + 7	—4.52	2.47 P<.05 >.02	—9.68 to + .64			

*Fisher *t* test for the significance of a mean difference.
**Fisher *t* test for the difference between two means, when samples are distinct.

TABLE 5

PER CENT OF CASES WHO DIFFERED ON THE KUHLMANN-ANDERSON FROM THE REVISED STANFORD BY THE LISTED RANGE OF *IQ* POINTS

Range of *CA*'s	Sub-class	No. of cases	Range of *IQ*'s on Stanford	Ranges of *IQ* points by which *K-A* differed from the Revised Stanford-Binet, Form *L*: —3 to +3	—4 and +4	—5 and +5	—26 to +21	Mean gain or loss in *IQ* points (*K-A*)-*S*
7-6	X_1	19	70 to 90	*73.7	*5.2	*10.6	*10.5	+ .35**
to	X_2	10	65 to 69	40.0	0.0	0.0	60.0	+3.70
13-1	X_3	33	46 to 64	24.2	9.1	15.2	51.5	+5.21
13-2	Y_1	14	70 to 100	28.6	0.0	14.3	57.1	—7.50
to	Y_2	9	63 to 69	44.4	11.1	11.1	33.4	+0.11
15-11	Y_3	27	45 to 62	29.6	7.4	7.4	55.6	+5.30

*These figures are to be read as percentages.

**These are to be read as *IQ* points.

markedly in their ability to discriminate between certain *IQ* ranges.

When we compare the Kuhlmann-Anderson *IQ*'s of those subjects in Groups X_3 and Y_3 (Table 5) with their Stanford *IQ*'s, we find that not only the *K-A* means are significantly higher than the Stanford,[7] but also that 11 of the 33 $K\text{-}A_{IQ}$'s now exceed the mean of the Stanford *X* group (Table 1) and that seven of the 27 $K\text{-}A_{IQ}$'s now exceed the mean of the Stanford *Y* group. These subjects had all scored below the *IQ* of 65 on the Stanford. Below these *IQ*'s, then, the Stanford discriminates between small *IQ* ranges to a better extent than does the Kuhlmann-Anderson.

Although the two tests differ markedly below the *IQ* of 65, there appears to be little difference between the two tests within the *IQ* range of 70 to 90 for the X_1 group. The mean gain in *IQ* points of the *K-A* over the Stanford is only +.35 points; and 89.5 per cent of the cases who scored between this range on the Stanford do not differ on the *K-A* by more than ±5 *IQ* points (Table 5). This compares well with the Terman and Merrill criteria which we have listed in Table 3.

Although the chances are very good that a child having an *IQ* between 70 and 90 on the Stanford (*CA*'s between 7-6 and 13-1) will maintain his score on the *K-A* within ±5 points, we cannot say, conversely, that one having an *IQ* on the *K-A* will have as much chance of maintaining his identical score on the Stanford. Thus, 17 of the 26 subjects who scored *IQ*'s of 70 to 91 on the *K-A* (*CA*'s between 7-6 and 13-1) differed from the Stanford by ±5 *IQ* points; and 8 of these 26 scored between the *IQ*'s of 57 and 64 on the Stanford. In other words, given an *IQ* of 70 on the *K-A*, our accuracy of prediction from the *K-A* to the Stanford diminishes.

Group Y_b in Table 4 does not follow the same pattern as Group X_b. Instead of showing a mean gain in *IQ* points on *K-A* over the Stanford, it shows a significant mean loss. The negative relationship between the *IQ* on the Stanford and *IQ* points gained by the *K-A*, evident in the X_a and X_b groups, here becomes more marked. Thus, given a Stanford *IQ* above 70 in sub-class Y_1, Table 5, we can be reasonably confident that the $K\text{-}A_{IQ}$ will be below 70 (87.5 per cent of the 14 cases will score from 1 to 26 points less on the *K-A* than on the Stanford). Within the *Y* groups, then,

[7]Since the *K-A* shows a significant mean gain in *IQ* points over the Stanford (Table 4), the difference between the mean *IQ*'s of each test is also significant.

the Terman-Merrill Test discriminates better at both ends of the distribution. There is also a tendency for the Stanford to discriminate to a better extent within the X_2 range. On the other hand, this tendency is very slight in the Y_2 group.

While a definite negative relationship exists between the Stanford *IQ*'s and *IQ* points gained by *K-A* over S.[8] the relationship between chronological age (when the *K-A* was given) and number of *IQ* points gained by *K-A* over the Stanford is small and insignificant. For the *X* group the Pearson *r* is —.007 with a *t* of .05; and for the *Y* group the *r* is —.27 with a *t* of 1.98.

Summary and Conclusions

The present research was designed to test two assumptions: First, that *IQ*'s on the Kuhlmann-Anderson, administered individually to mental defectives, are comparable to *IQ*'s on the Revised Stanford Binet over a wide range of chronological ages; and second, that both tests will discriminate to the same extent between different *IQ* ranges.

To test these hypotheses we carried out the following experiment: From the total population at the Minnesota School and Colony between the ages of 7-6 and 15-11, we eliminated all children with speech, sensory, or physical defects and reading disabilities. We selected 112 children at random from those who remained. The Kuhlmann-Anderson *Intelligence Tests* and the Revised Stanford-Binet were administered individually to these children by two experienced examiners. Since we used the method which each author recommends for computing an *IQ*, we had to divide our total group of 112 into two sub-groups: those with *CA*'s between 7-6 and 13-1 and those between 13-2 and 15-11. An analysis of our data yielded the following results:

1. A significant difference, in favor of the Stanford, existed between the mean *IQ*'s and mental ages in the 7-6 to 13-1 group.

2. The mean differences in *IQ*'s and *MA*'s in the 13-2 to 15-11 group was in favor of the Stanford but was not significant. This did not, however, mean that the *IQ*'s of the two tests were comparable.

3. When we compared the deviations in *IQ* points of the *K-A*

[8]For the *X* group, Table 1, the Pearson *r* is —.41 with a *t* of 3.55 which is significant at the 1 per cent level. For the *Y* group the *r* is —.74 with a *t* of 7.81 which is significant at the 1 per cent level.

from the Stanford, we found that only those subjects who scored *IQ*'s between 70 and 90 on the Stanford, in the 7-6 and 13-1 group, would score comparable *IQ*'s on the *K-A*.

4. Our data also showed that the *K-A* does not discriminate between small *IQ* ranges as well as the Stanford. This failure to discriminate was especially marked for the 7-6 to 13-1 group below the *IQ* of 64 (Stanford) and for the 13-2 to 15-11 group below the *IQ* of 62 and above the *IQ* of 70.

5. The relationship between chronological age at the time of the *K-A* test and number of *IQ* points gained by *K-A* over the Stanford was negative, small, and insignificant.

We should therefore, suggest that experimenters do not consider scores on the *K-A* to be equal to the same score on the Stanford. When fine discriminations between mental ranges of a population such as ours are desired, the Stanford rather than the Kuhlmann-Anderson should be employed. Further, it is doubtful whether the Kuhlmann-Anderson should be employed as an individual examination when diagnosing a case of mental deficiency.

References

1. Carlton, T. A comparison of the Revised Stanford-Binet, Form *L*, with the Kuhlman Tests of Mental Development: I. Mentally defective adolescents. (To be published.)
2. ———. A comparison of the Revised Stanford-Binet, Form *L*, with the Kuhlmann Tests of Mental Development: II. Performances of mentally defective children. (To be published.)
3. Fisher, R. A. Statistical Methods for Research Workers. London: Oliver & Boyd, 1936.
4. Goulden, C. H. Methods of Statistical Analysis. New York: Wiley, 1939.
5. Gray, W. S. Standardized Oral Reading Paragraphs Test. Bloomington, Ill.: Public School Pub., 1939.
6. Hartman, A. A. A comparison of the ratings of mentally retarded prisoners on Army Alpha and Standard-Binet. *J. App. Psychol.*, 1939, **23**, 608-611.
7. Hilden, A. H., & Skeels, H. M. Comparison of the Stanford-Binet Scale, the Kuhlmann-Anderson Group Test, the Arthur Point Scale of Performance Tests, and the Unit Scales of Attainment. *J. Exper. Educ.*, 1935, **4**, 214-230.
8. Kefauver, G. N. Need of equating intelligence quotients obtained from group tests. *J. Appl. Psychol.*, 1929, **19**, 92-101.

9. KREEZER, G. L., & BRADWAY, K. P. The direct determination of the probable error of measurement of Binet mental age. *J. Educ. Res.*, 1939, **33**, 197-214.

10. KUHLMANN, F. The Kuhlmann-Anderson Tests compared with seven others. *J. Appl. Psychol.*, 1928, **12**, 545-594.

11. ———. Tests of Mental Development. Minneapolis: Educational Test Bureau, 1939.

12. KUHLMANN, F., & ANDERSON, R. G. Intelligence Tests for Ages Six to Maturity: Instruction Manual. Minneapolis, Minn.: Educational Tests Bureau, 1927.

13. LINDQUIST, E. F. Statistical Analysis in Educational Research. New York: Houghton Mifflin, 1940.

14. MERRILL, M. A. Significance of *IQ's* on the Revised Stanford-Binet Scales. *J. Educ. Psychol.*, 1938, **29**, 641-651.

15. MILLER, W. S. The variation and significance of intelligence quotients obtained from group tests. *J. Educ. Psychol.*, 1924, **15**, 359-366.

16. TERMAN, L. M., & MERRILL, M. A. Measuring Intelligence. New York: Houghton Mifflin, 1937.

Child Guidance Clinic
Sunbeam Home Association
Oklahoma City, Oklahoma

The Journal of Genetic Psychology, 1942, 60, 99-108.

WHAT ADOLESCENT AND ADULT STAMP COLLECTORS LEARN FROM THEIR AVOCATION*

Department of Psychology, Clark University

DONALD E. SUPER AND RALPH L. CARLSON

In previous publications (2, 3) the senior author has questioned certain widely held opinions concerning avocations. In one of these he presented data on adolescent stamp collectors which led to the conclusion that high school stamp collectors are not better informed in the field of social studies than are non-collectors of the same age and intelligence, at least insofar as socially significant information is concerned. A comparison of strictly factual knowledge was also made, on the hypothesis that stamp collectors might learn more about the location and size of towns and countries, languages, dynasties, and currencies than non-collectors. The available data were not sufficient to make possible adequate conclusions on this point, although it appeared that the collectors again had no advantage over the non-collectors.

A. PROBLEM

The purpose of the studies reported in this paper was to determine whether adolescent stamp collectors, who do not acquire any socially significant information concerning world affairs as a result of their collecting activities, learn more than non-collectors concerning strictly factual historical and geographical matters, and whether the facts concerning adolescent collectors apply also to adults.

B. METHOD: ADOLESCENTS

As a check on the findings of the previous study, it was repeated on a different sample of high school students; supplementary techniques were used to provide the data needed to answer the newly raised questions. The procedures were as follows.

A total of 119 juniors and seniors, of both sexes, were tested in the Fall of 1939 in connection with the Clark University Guidance Service. These students were selected by the principals of 15 high

*Received in the Editorial Office on October 3, 1940.

schools in Worcester County, Massachusetts, as individuals in need of vocational and educational guidance, and were tested by the writer in two sessions, all students being present each time. The tests with which we are concerned in this paper are the Otis *Self-Administering Test of Mental Ability, Higher Form A;* the *Co-operative General Achievement Test, Form O, Part I, Survey of the Social Studies;* a *Test of World Knowledge*, and a *Test for Stamp Collectors*, both of which were devised by the writer for the purposes of this study.

The *Test of World Knowledge* took 15 minutes and consisted of 119 items of general information which might be learned directly from postage stamps or from a study of their background. Typical questions are:

> The Karageorgevitch dynasty is: (1) Russian, (2) Serbian, (3) Croat, (4) Polish, (5) Bulgarian.
>
> The founder of the Chinese Republic was named: (1) Sun Yat-Sen, (2) Hu Shi, (3) Chiang Kai-Shek, (4) Yin Yang, (5) T. V. Soong.
>
> Macao is in ———. Honkong belongs to ———.
>
> The Charter Oak is in: (1) Massachusetts, (2) Connecticut, (3) Virginia, (4) Rhode Island, (5) New York.
>
> Balboa discovered the Pacific in: (1) 1492, (2) 1498, (3) 1502, (4) 1513, (5) 1525.

These and similar items were selected by scrutinizing American and foreign postage stamps for historical and similar facts included in their pictures and inscriptions, and by reading articles in back numbers of various stamp magazines and catalogues. The items for the *Test for Stamp Collectors* were chosen in the same manner and from similar material, plus a booklet on stamp collecting (1) which included a brief history of collecting and facts about famous stamps and collectors. The purpose of this last test was to test knowledge of the methods, principles, and history of stamp collecting. The time limit for this 50-item test was five minutes. Sample items follow.

> *Stamps* is published in: (1) Buffalo, (2) New York City, (3) Louisville, (4) Houston, (5) Topeka.
>
> The first postage stamp was issued in: (1) England, (2) France, (3) Germany, (4) Italy, (5) United States.
>
> Wove is a kind of: (1) watermark, (2) color, (3) design, (4) cloth, (5) paper.

The greatest stamp collection ever assembled belonged to: (1) Wanamaker, (2) George V, (3) Roosevelt, (4) Schwab, (5) Ferrari.

Each student also indicated whether or not he had ever collected stamps, whether he was still an active collector, and the years in which he had begun and ceased collecting stamps.

C. Findings: Adolescents

Of the total group of 119 students, 65 had at one time collected stamps and 54 had never collected stamps. As the former group had all collected until at least a few years before the study was made, and for a period of more than a year, these two groups were first compared to each other. The results appear in Table 1.

TABLE 1
Intelligence and General Knowledge of One-Time Collectors and Non-Collectors

Test	Collectors' mean	Non-Collectors' mean	D/σ_D
Intelligence (*IQ*)	113.90	113.50	.25
Social Studies	59.70	57.50	1.40
World Knowledge	24.90	22.10	1.26
Stamp Collecting	6.60	3.74	3.25
Number	65	54	

The two groups are seen, in Table 1, to be of equal intelligence, the Otis *IQ*'s being 113.9 for those who once collected and 113.5 for those who never collected, and the critical ratio, .25. There is a slight difference in the test of significant social knowledge, the collectors' mean standard score being 59.7 and the non-collectors' being 57.5, but the difference is not significant, the ratio being 1.40. The situation in this group is, therefore, approximately the same as that in the first study (3). The scores on the test of factual information such as might be gleaned from stamps also show no significant difference between the groups, although the ratio of 1.26 is in favor of those who then or at one time collected stamps. The results of the final comparison, that involving knowledge of stamp collecting, are important in that they show that the collectors have learned something about stamp collecting, this group making a mean raw score of 6.6, which is significantly higher than the non-collectors' mean of 3.7 since the critical ratio is 3.25.

This last point suggests that the procedure of grouping recent and present stamp collectors together is justifiable in this study, for if the one-time collectors have learned more in one area regardless of how recently they have collected, what they have learned in other areas should also have been retained. It seemed wise, however, also to compare those who still collected with those who had never collected, as late middle adolescents might collect differently from early middle adolescents. Only 12 students were active collectors at the time of the study; these are compared in Table 2 with the 54 who had never collected.

TABLE 2
INTELLIGENCE AND GENERAL KNOWLEDGE OF COLLECTORS AND NON-COLLECTORS

Test	Collectors' mean	Non-Collectors' mean	D/σ_D
Intelligence (*IQ*)	114.00	113.50	.28
Social Studies	58.66	57.50	.80
World Knowledge	35.38	22.10	7.81
Stamp Collecting	8.35	3.74	3.29
Number	12	54	

Again there are no differences in mental ability ($CR = .28$) and significant social knowledge ($CR = .80$), and again the stamp collectors are better informed on the history and methods of stamp collecting than the non-collectors ($CR = 3.29$). But these active adolescent stamp collectors have learned facts of history and geography which are not known by the non-collectors, the mean scores being 35.38 and 22.1 and the critical ratio, 7.81.

D. PROBLEM: ADULTS

It seemed likely that differences between the collecting habits of adolescents and adults might result in different findings for the last-named group. Adult collectors are known to be rather more systematic in their collecting: perhaps they not only acquire more factual information in the field of the social studies but also a greater understanding of significant social problems than adults who do not collect stamps. Since important differences in the avocational activities of adolescents and adults had already been found in a study concerned with interest patterns (2), it seemed desirable to investigate adult collecting with techniques similar to those used in the adolescent studies.

E. METHOD: ADULTS

The methods and instruments used were the same, insofar as possible, as those used in the present study of adolescent collectors. The instruments were identical: the Otis *Self-Administering Test of Mental Ability, Higher Form A* (20-minute time limit); the *Coöperative General Achievement Test, Form O, Part I, Survey Test of the Social Studies;* the Super *Test of World Knowledge;* and the Super *Test for Stamp Collectors.* The last two instruments, previously described, were found to have corrected odd-even reliability coefficients of .91 and .80, respectively, sufficiently high for group comparisons. With this battery it was possible to measure mental ability, significant social knowledge, factual historical and geographical information which might be gathered from postage stamps, and knowledge of stamp collecting. This distinction between the two types of achievement in the social studies is justified by a correlation of .48 between the two tests, or .16 with intelligence partialled out, showing that although they do have something in common, other factors are more important.

The coöperation of one of the two local city-wide stamp clubs was obtained, and the tests were administered to its members at regular meetings. Although the other local stamp club did not coöperate officially, many of its members were tested because of overlapping membership and because an effort was made to test all other adult stamp collectors who could be located in Worcester. These persons were located by asking club members to suggest other collectors who might be asked to coöperate and through various miscellaneous sources. There is no way of knowing how complete or typical of adult stamp collectors the sample is; as will be seen, however, the subjects come from a variety of occupational levels and, as the hobby is highly organized, it seems likely that the group is fairly representative.

The finding of adequate controls presented something of a problem, for there were no obvious and easily available control groups as in the adolescent studies. The adult collectors had been selected from the community as a whole because they collected stamps; control subjects, non-collectors, had perforce to be selected from the same community in such a way as to make them as comparable as possible. The difficulty was increased by the nature of the study, for although it was relatively easy to get the coöperation of stamp collectors, who were naturally interested in a study of stamp col-

lecting, there was no such effective appeal to motivate other adult groups or individuals. The junior author, therefore, obtained the coöperation of church and Y. M. C. A. groups, their friends, and some of his friends, attempting to get subjects similar to the collectors in socio-economic status. After a fairly large number of such controls or non-collectors had been tested, the two groups were compared in terms of socio-economic status, mental ability, and age, and cases were discarded until a fair balance was achieved.

In all, 39 collectors and 56 non-collectors were tested. Of these, 7 collectors (3 students, 4 semi-skilled) and 16 non-collectors (6 students, 10 professional) were discarded. The validity of the division into collectors and non-collectors is indicated by the fact that the former had collected stamps for 10 years in the average, the latter not at all, according to their own statements, and by the critical ratio (D/σ_D) of 10.22 obtained for the difference between their mean scores on the test of stamp-collecting information.

Table 3 gives the socio-economic classifications of the two groups

TABLE 3

SOCIO-ECONOMIC LEVELS OF COLLECTORS AND NON-COLLECTORS

Occupational level	Collectors N	Collectors %	Non-Collectors N	Non-Collectors %
1. Professional	12	37.5	16	40.0
2. Proprietary-Manager	4	12.5	1	2.5
3. Commercial	9	28.1	15	37.5
4. Skilled	5	15.6	8	20.0
5. Semi-skilled	2	6.3	0	—
6. Unskilled	0	—	0	—
Mean rank	2.4		2.4	

used in this study. The proportions of professional people in each group are about equal, 37.5 per cent and 40 per cent. There are rather more proprietors and managers in the collecting group, 12.5 per cent as opposed to 2.5 per cent. On the other hand, there are more non-collectors from the commercial level, 37.5 per cent versus 28.1 per cent, and more from the skilled level, 20 per cent versus 15.6 per cent. Slightly over 6 per cent of the collectors are from the semi-skilled level, none of the non-collectors. No subjects came from unskilled occupations, which, as the collectors are presumably otherwise unselected, suggests, as did another study (2), that stamp collecting as a hobby is restricted to middle and higher socio-economic levels.

The mean rank of the collecting and non-collecting groups in each case is 2.4, which indicates that the differences at each socio-economic level tend to counterbalance each other, making the two groups quite comparable in this respect.

In Table 4 it is made clear that the experimental and control

TABLE 4
INTELLIGENCE OF EXPERIMENTAL AND CONTROL GROUPS

	N	Mean IQ	SD	D/σ_D
Collectors	32	118	14.4	.29
Non-collectors	40	117	14.5	

groups are quite comparable as to intelligence, the mean IQ's being 118 and 117 and the critical ratio, .29.

In Table 5 are presented the data on the ages of the two groups,

TABLE 5
AGES OF EXPERIMENTAL AND CONTROL GROUPS

	N	Mean age	SD	$D\sigma/_D$
Collectors	32	41	13.2	2.16
Non-collectors	40	35	9.7	

the collectors being older by six years than the non-collectors. We did not succeed very well in matching our two groups in this respect. However, the difference is not statistically significant, although it begins to approach significance, the critical ratio being 2.16. As 68 per cent of the collectors are between 28 and 54 years old, and a similar majority of non-collectors are between 25 and 45 years old, it is also questionable as to whether the difference is psychologically significant, for the two groups reached maturity at about the same time and, therefore, have had essentially similar experience insofar as exposure to cultural influences are affected by chronology.

From these three comparisons it is concluded that, since the socio-economic level, mental ability, and age of the two groups are essentially similar, the general cultural influences to which they have been exposed, and their ability to be affected by these, are roughly the same. Presumably, then any differences found in the other tests administered are the result of a specific cultural influence. As the controlled, differing variable is stamp collecting, and as other uncon-

trolled variables may fairly be assumed to be randomly distributed with the controlled, equated variables, differences in the social information and understanding of the collectors and non-collectors may be attributed to factors associated with stamp collecting as an avocation.

F. Findings: Adults

The means and critical ratios of the scores on the achievement tests are presented in Table 6.

TABLE 6
Social Information and Understanding of Stamp Collectors and Non-Collectors

	Collectors' mean score	Non-collectors' mean score	D/σ_D
Stamp collecting	25	6	10.22
Factual world knowledge	49	27	5.02
Significant social knowledge	68	64	1.42

The first row shows that the stamp collectors, as has already been mentioned in another connection, know more about stamp collecting as a hobby than the non-collectors, the mean scores being 25 and 6, the critical ratio 10.22. The second row reveals a real difference in general knowledge which might be acquired through collecting, the collectors making a raw score of 49 on the test of factual world knowledge, the non-collectors 27, the critical ratio being 5.02. In the third row it appears that there is no real difference in the understanding of significant social problems shown by the two groups, the mean scores being 68 and 64, the critical ratio, 1.42.

Further light is thrown on the principal question of this study by correlating the number of years of active stamp collecting (for the collecting group only) with scores on the achievement tests. These figures are given in Table 7.

TABLE 7
Relationship between Collecting and Achievement

	r: No. of years a collector
Significance social knowledge	.07
Factual world knowledge	.25
Stamp collecting information	.50

In this table it appears that the number of years during which one has collected stamps has no relationship to significant social understanding, is positively but not closely associated with factual information of a geographical or historical type, and rather closely related to mastery of the methods and history of stamp collecting.

G. Discussion

The adolescent study has clarified the issues left doubtful by its precursor in that it has shown that high school students who are active stamp collectors do acquire more information of a factual historical and geographical type than non-collectors, together with more information about the history and facts of stamp collecting, even though those who had at one time collected stamps were better informed only in the latter area. This last fact suggests that late middle adolescents (the active collectors) collect stamps in a manner different from that typical of early and middle adolescents (this is what the former collectors were when they collected), studying the stamps more carefully and learning more about them and their background. Adult collectors, like older adolescent collectors, know a good deal more about stamp collecting, something more about the factual aspects of geography and history, but nothing more about more socially significant social problems, than non-collectors.

The implications of these findings for educational practice are, briefly, as follows. First, it is clearly not safe to assume that extra-curricular and avocational activities have a given educational value because of superficial resemblances to other educational activities: these values and outcomes must be demonstrated. This finding has been confirmed by the senior author in other ways (2). Secondly, stamp collecting, when systematically pursued as by adult collectors, is associated with the learning of historical and geographical facts and can thus be used by teachers of the social studies as an instructional technique, even though it does not contribute to social understanding.

H. Summary

Tests of significant social knowledge, of factual historical and geographical information, of information about stamp collecting, and of intelligence were administered to 65 adolescent and 39 adult stamp collectors. Fifty-four adolescent and 56 adult non-collectors were also tested. Adult subjects were dropped in order roughly to

equate collectors and non-collectors for socio-economic status, age, and mental ability. This left 32 stamp collectors and 40 non-collectors in the adult control and experimental groups. The adolescent groups were equal as a result of the sampling method. A comparison of the mean scores of the above tests of achievement, and an analysis of their intercorrelations, showed that:

1. Stamp collecting among adolescents and adults is associated with increased knowledge of the history and principles of stamp collecting.

2. Stamp collecting among adolescents and adults is not associated with an increase in significant social information and understanding.

3. Stamp collecting among adults is associated with the acquisition of a greater amount of information concerning the facts of history and geography; among adolescents, only if they have been active during the last years of high school.

4. Compared to the general population, adult stamp collectors are of superior mental ability and, to a lesser extent, socio-economic status; adolescent collectors resemble other high school students (already a selected group).

5. The distinction between significant social knowledge and factual world knowledge is valid.

Implications for educational practice were briefly pointed out, together with the need for more objective evaluation of other avocational and extra-curricular activities.

References

1. Renouf, A. Stamp Collecting. New York: Leisure League, 1934.
2. Super, D. E. Avocational interest patterns: A study in the psychology of avocations. Stanford University: Stanford Univ. Press, 1940.
3. ———. The educational value of stamp collecting. *J. Educ. Psychol.* 1940, **31**, 68-70.

Department of Psychology
Clark University
Worcester, Massachusetts

The Journal of Genetic Psychology, 1942, 60, 109-119.

STUDIES IN VISION: II. DARK ADAPTATION IN NORMAL AND A-AVITAMINOTIC RATS*[1]

Laboratory of Physiological Psychology, Harvard University

CLIFFORD T. MORGAN

A. PROBLEMS

This paper reports a reasonably successful attempt to obtain curves of dark adaptation in the rat under normal and hypovitaminotic conditions. The study was begun with three objectives in mind. The first of these was the quantitative determination of the rôle of vitamin-*A* in visual processes as measured by the behavioral reactions of an animal to light. The rat was chosen as subject because methods of vitamin-*A* depletion have been best standardized on this animal and because the experimental methods required by the problem are readily applied to the rat.

Other researches on animals have shown that the visual sensitivity of vitamin-*A* depleted subjects is impaired by the vitamin-deprivation, but quantitative statements of the relationship between vitamin-*A* and visual sensitivity have been lacking in these studies (5, 15, 29, 30). In recent years, it has often been reported by clinicians that nightblindness (hemeralopia) may be ameliorated by the administration of vitamin-*A*, but such studies have lacked the thoroughness required to reveal fundamental facts concerning underlying visual processes (2, 4, 9, 16, 17, 18, 28).

After the present study had in large part been completed, two researches, similar to the present one except that man was used as subject, were published; one by Wald, Jeghers, and Arminio (33)

*Received in the Editorial Office on October 3, 1940.

[1]This paper is the substance of part of a dissertation presented to the Department of Psychology, University of Rochester, in partial fulfillment of the requirements for the degree Doctor of Philosophy (1939). The work was carried out partly under Professor Leonard Carmichael and partly under Professor Elmer Culler. Special acknowledgment is due Dr. A. B. McCoord and Dr. Samuel Clausen, University of Rochester School of Medicine, for their interest and labors in behalf of the research. They prepared the vitamin-*A* deficient animals used in the study and rendered invaluable suggestions toward the solution of problems which arose. The kind assistance of Professor Karl U. Smith, University of Rochester, is also gratefully acknowledged.

and another by Hecht and Mandelbaum (14). The material presented in these papers is in many respects superior to that summarized here, yet our data differ in certain details and throw some additional light on the problem.

It was a further purpose of the present study to test the adequacy of Hecht's photochemical theory (11, 12) on some vertebrate animal other than man, particularly one whose eye was atypical with respect to the character of the retinal elements. The rat is such an animal. So far, all dark-adaptation data which are extensive enough for mathematical treatment according to Hecht's theory have been obtained either from man (10, 13, 32) or from invertebrate animals (see 11, 12, 13).

The third purpose of the study was to supply information—possibly crucial information—concerning the question of duplex vision in the rat. Most of the past researches have attacked the problem by attempting to determine whether or not the rat can discriminate color *per se*. Watson and Watson (36), Munn (24, 25), and Coleman and Hamilton (1) failed to obtain discrimination of color when all other properties of the visual stimulus were controlled. On the other hand, Walton (35), Muenzinger and Reynolds (27), and Munn and Collins (26) claim that the discrimination of color is possible in the rat if training is continued long enough under the proper conditions.

When the problem of duplex vision was approached via the luminosity curve and electrical methods of recording retinal potentials, Graham and Riggs (7) secured negative results. It may be questioned, however, whether the electrical method of determining the visibility curve is at all satisfactory for photopic levels of illumination (see Graham, Kemp, and Riggs, 6). The results of Graham and Riggs therefore probably have little significance for the problem of color vision in the rat.

Anatomic evidence is no less equivocal than the electrical and behavioral data. For a long time, it was believed on good histological evidence (see 19) that the eye of the rat possessed rods exclusively, but recently Menner (21) and Walls (34) have argued that cones may be found in the rat's retina when certain histological methods of preparation as well as particular structural differentia are employed.

A summary of the evidence for and against rod-cone vision in the rat thus leaves one uncertain as to the conclusions which may be

drawn. Now in the case of man, the dark-adaptation curve consists of two distinct segments which are presumably correlated with the respective functions of the rods and cones. If the separation of the human dark-adaptation curve into two parts is really an indication of duplicity in retinal function, one might well attempt to settle the question of duplex vision in the rat from dark-adaptation data. This then was the third object of our study.

B. Apparatus and Method

The present study involved, in its preliminary phases, the development and construction of a new discrimination apparatus. This device has been fully described in an earlier paper (23) and need not be discussed here.

Animals were always required to discriminate a dark from a lighted panel. They signified their choice by depressing a lever located beneath each of the two panels. When the choice was correct, a "shot" of food was delivered into a tray situated midway between the two panels (8, 23). The animals were shocked for an incorrect choice.

Rats were trained with the brightness of the lighted panel at 3.2 millilamberts. About 150 to 250 trials were required to reach the criterion of 38 correct choices in 40 successive trials. This was exclusive of 50 preliminary trials in which all lever-responses were rewarded. After the original light-dark discrimination had been mastered, the brightness of the illuminated panel was reduced step by step until the lower limit of discriminability was reached.

In dark-adaptation tests which were begun after the preliminary training just described was completed, each daily session with a rat involved a "pre-test" and a "test" procedure. Pre-test trials were controls to make certain that a rat was fully capable of passing at the brightness level at which he was to be tested in dark-adaptation trials. Accordingly, the rats were required in the pre-test period to give either 10 correct responses in 10 trials or 18 correct responses in 20 trials.

After pre-test trials, the animals were subjected to light-adaptation for one hour. This was done in a box, 8 by 10 by 12 inches, whose inner surfaces were painted a glossy white and illuminated by means of a 150-watt lamp which was located on the rear wall of the box. The light-adaptation period was one hour in length in order to insure a high and rather constant degree of light-adapta-

tion in all tests and in all animals. Rats were always quite active during this period; they never seemed to sleep nor did they appear to be attempting to avoid the light.

Following light-adaptation, dark-adaptation tests were given at the brightness level used in the pre-test period, and this one brightness was the only one used in the dark-adaptation tests of one day. The brightness was used on successive days, usually three to seven days, until the "threshold-time" of that brightness could be ascertained.

It was, of course, threshold-time rather than threshold-brightness which was the dependent variable in our procedure, for the brightness-level was always kept constant throughout any period of testing. The rat simply made lever-responses as rapidly as it ate and was disposed to press, and the dark-adaptation time of each response was recorded. Correct and incorrect responses, recorded in this way, were summed for each five-minute period of dark-adaptation. Daily tests at one brightness were continued until about 50 responses had been obtained for each five-minute period. The percentage of correct responses for each period was then computed, and the threshold-time assigned by interpolation of the 75 per cent point.

Rats were ordinarily tested for one-half hour. At high brightness-levels, where the threshold-time was rather small, shorter intervals were employed. The time of dark-adaptation at which testing was begun for a given test-brightness was found by several days of trial-and-error and by reference to thresholds previously secured. In the case of every animal, several "practice-tests" were given the animal before taking threshold-data in order to accustom the rat to the light-adaptation box and also in order that the experimenter might determine the proper time of dark-adaptation at which to begin testing.

The motivation of the animals was controlled by the rationing of food to individual rats.

C. Results

In accordance with established procedure used in the preparation of vitamin-*A* deficient rats, 27 female albino rats taken from four litters were weaned at 20 to 23 days of age and placed at that time on a vitamin-*A* free diet. About six weeks on such a diet were sufficient for the appearance of the usual symptoms of vitamin-*A* deficiency. After this period training and testing were begun.

The time required to make the desired measurements was much longer than that during which depleted animals ordinarily live. After some experimentation, it was found that the oral administration of about four units of the vitamin was capable of producing growth in animals whose weights were falling off rapidly. The effects of the dose usually lasted from 8 to 15 days. The administration was repeated whenever declining weight of an animal indicated that death from vitamin-*A* deficiency was near. In spite of these precautions, eight rats of an initial group of 27 died before they could be trained, and nine more succumbed before thresholds could be obtained. Threshold-data were finally secured upon 10 female hypovitaminized rats.

A group of four female rats was used to obtain normal data with which to compare vitamin-*A* deficient animals. The normal group was maintained on exactly the same *A*-free diet except that weekly

TABLE 1

THRESHOLDS DURING DARK-ADAPTATION OF HYPOVITAMINIZED RATS (TIME IN MINUTES)

	Log Micromillilamberts								
Subject	6.5	6.0	5.5	5.0	4.5	4.0	3.5	3.0	2.5
35	6.4 (3)		10.3 (8)		18.8 (4)	21.1 (8)		53.9 (4)	
38		6.3 (3)		18.5 (6)					
41	2.9 (4)	4.3 (4)		12.2 (8)		27.5 (12)		43.3 (8)	76.6 (3)
42	4.4 (3)				17.8 (2)				58.3 (8)
44		8.9 (9)	9.3 (4)		21.8 (7)	26.5 (7)		56.6 (5)	81.0 (4)
48	12.7 (4)	20.9 (8)					32.9 (3)		87.0 (8)
57		11.7 (5)			17.1 (4)		31.4 (8)		
58							36.3 (3)		
60				18.9 (9)		37.3 (4)			
62		5.1 (4)	13.7 (9)			21.2 (5)		42.7 (8)	51.1 (12)

The figures in parentheses indicate the number of days after the administration of a minimal dose of viatmin-*A* at which the threshold tests were begun.

TABLE 2
THRESHOLDS DURING DARK-ADAPTATION OF NORMAL RATS
(TIME IN MINUTES)

Log micro-millilamberts	Subjects 1	12	20	30
4.0	5.1	5.1	4.6	3.5
3.5	11.5	9.7	9.9	8.2
3.0	16.5	14.9	13.5	13.7
2.5	19.5	23.2	20.5	
2.0	27.3	31.0	29.4	25.4
1.5	38.4	46.2		39.3

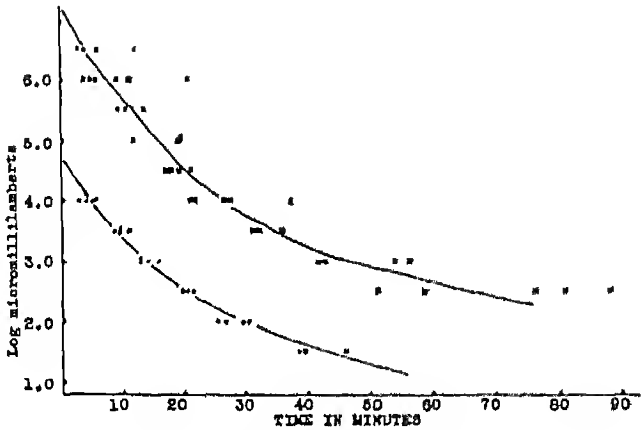

FIGURE 1
THE DARK-ADAPTATION CURVES OF NORMAL (DOTS) AND VITAMIN-*A* DEFICIENT (SQUARES) RATS

doses of cod liver oil were administered. The vitamin-*A* depleted and normal animals were of approximately the same age.

Tables 1 and 2 and Figure 1 summarize the time-thresholds of dark-adaptation in normal and *A*-avitaminotic rats. Table 1 includes the thresholds secured for the 10 hypovitaminized animals; comparable values for the four normal animals appear in Table 2. The brightness range of the tests is, as indicated, 3.2 x 10^6 micromillilamberts to 3.2 x 10^1 micromillilamberts.

From these data, the following observations may be made: (*a*) Dark-adaptation of the normal rat proceeds relatively slowly, for it

is still definitely in progress at the end of 50 minutes. (*b*) The curve of the *A*-avitaminotic subjects is displaced about two log units above that of the normal rat, but otherwise shows the same general form as that of the normal subjects. (*c*) There seems to be no shift in the adaptation-curve of the rat which could be interpreted as representing a division of rod- and cone-function such as is typically found in similar curves for man. (*d*) The administration of minimal sustaining doses of vitamin-*A* apparently has no significant relation to the thresholds of the *A*-deficient animals. This is indicated by the fact that no consistent variation in threshold values occurs with varying periods after administration of sustaining doses of vitamin-*A*. (*e*) The relative consistency of the thresholds of hypovitaminosis indicates the remarkable degree of homogeneity of the group of rats with respect to the visual defect which results from the degree of hypovitaminosis obtained in the present study.

D. Discussion

In the introduction to this paper three problems were raised: (*a*) the question of whether or not the importance of vitamin-*A* in the visual process that has been claimed in clinical work can be substantiated in a crucial experiment in which unrestricted control of vitamin-*A* storage is possible; (*b*) the nature of the dark-adaptation process in the rat and the conformity of this process to the photochemical theory of vision; (*c*) the question of duplicity of function in the visual mechanism of the rat. Let us discuss these problems in the order mentioned.

The results obtained here indicate without question the importance of vitamin-*A* in processes of vision. They do not necessarily indicate that vitamin-*A* is involved in the photoreceptor process itself, for it is conceivable that the effects obtained might have resulted from secondary relations of vitamin-*A* chemistry to the photochemical process. Wald's finding (31), however, that vitamin-*A* is liberated in the decomposition of visual purple, seems to indicate that the substance is of primary importance in vision.

Although Wald, Jeghers, and Arminio (33) and Hecht and Mandelbaum (14) have reported similar results for man, the present findings clear up a question not satisfactorily answered by the other studies. One cannot obtain nightblindness in all human individuals placed on a vitamin-*A* free diet; only a few, particularly those suffering from liver dysfunction (see 14), will react. This

is probably related to the fact that the storage of vitamin-*A* in the human individual is exceedingly difficult to deplete. The nightblindness produced in rats when standard conditions of vitamin-*A* depletion are observed is exceedingly uniform and can leave no doubt as to the genuine nature of the effect.

It should be pointed out, however, that the nightblindness obtained in rats was by no means complete. Yet one might expect more pronounced defects because the vitamin-*A* deficiency was undoubtedly acute. The meaning of this is not clear, but it might be argued that vitamin-*A* is not concerned as fundamentally in the visual process as the physiological results of Wald (31) would lead one to believe.

If the present curve of dark-adaptation in *A*-avitaminosis is compared with similar curves of the human studies, it will be noticed that the curve above shows a much more marked deviation from the normal curve in its initial portion than do the latter curves. No explanation of this difference seems particularly attractive at present. It may be a genuine difference between the effects of vitamin-*A* depletion in man and in the rat. Then again it may be due to some alteration in the behavior of the rats in the light-adaptation box which permitted greater adaptation to the light in depleted animals than in normal rats.

The question of the nature of the dark-adaptation process in the rat may now be discussed. Probably the first important point to be mentioned is that the rat's curve has less slope, that is, it descends more slowly and for a longer time, than the human dark-adaptation curve. An examination of the data from which the curve was computed indicates that this fact may be accounted for, at least in part, on the basis of the threshold-criterion which was used. (The reader will recall that 75 per cent passing was taken as the threshold-time.) The curves of improvement in performance from which the thresholds are computed are found to have less slope for the lower brightness values. This means that had a lower criterion of passing been used, the dark-adaptation curve itself would seem to possess a velocity which is more nearly that of the normal human curve.

The bearing of the present data on Hecht's photochemical formulations is unfortunately clouded by insufficient accuracy of the observations themselves. The dark-adaptation curves of individual animals and of the group are fitted fairly well by the bimolecular

equation (see 12), but certainly no crucial significance can be attached to the fit.

It is similarly impossible to reach, from these data, any assured conclusions concerning duality of visual function. It is certainly true that two segments, such as one expects to find in the eye possessing rods and cones, are not evident in the dark-adaptation curve obtained here. This cannot be discounted, but like the other studies already reported on the rat, it does not constitute crucial proof on the question, for it may be argued that not enough threshold values have been obtained and that the curve quite possibly does not cover the range in which cones function, if they do, in the rat. It is also possible, granting duplex vision in the rat, that rod- and cone-processes are not as clearly segregated functionally in the rat as they are in man, and that it therefore may be more difficult to demonstrate experimentally relationships dependent upon them.

E. Summary

The dark-adaptation of albino rats was measured by means of a discrimination-technique in which rats were required to choose between a lighted and a dark panel. The dark-adaptation curve of four normal and 10 *A*-avitaminotic rats was secured. It was found that: (*a*) the curve of hypovitaminotic subjects was displaced about two log units above that of the normal rat, but it was otherwise of the same general form as that of normal subjects; (*b*) there was no shift in the adaptation-curve of the rat which could be interpreted as representing a division of rod- and cone-function such as is typically found in similar curves of man.

References

1. Coleman, T. B., & Hamilton, W. F. Color blindness in the rat. *J. Comp. Psychol.*, 1933, **15**, 177-181.
2. Corlette, M. B., Youmans, J. B., Frank, H., & Corlette, M. G. Photometric studies of visual adaptation in relation to mild vitamin-*A* deficiency in adults. *Amer. J. Med. Sci.*, 1938, **195**, 54-65.
3. Crozier, W. J., & Wolf, E. Dark-adaptation in Agriolimax. *J. Gen. Physiol.*, 1929, **12**, 83-109.
4. Frandsen, H. Investigation of hemeralopia as evidence of minor degrees of vitamin-*A* deficiency. *Nutr. Abst. Rev.*, 1935, **4**, 621.
5. Fridericia, L. S., & Holm, E. Experimental contribution to the study of the relation between nightblindness and malnutrition. *Amer. J. Physiol.*, 1925, **73**, 63-77.

6. Graham, C. H., Kemp, E. H., & Riggs, L. A. An analysis of the electrical retinal responses of a color-discriminating eye to lights of different wave-lengths. *J. Gen. Psychol.*, 1935, **13**, 275-296.

7. Graham, C. H., & Riggs, L. A. The visibility curve of the white rat as determined by the electrical retinal response to light of different wave-lengths. *J. Gen. Psychol.*, 1935, **12**, 279-295.

8. Hanford, H. M., & Morgan, C. T. A simplified lever-operated device for conditioning studies in rats. *J. Gen. Psychol.*, 1939, **21**, 229-232.

9. Haurowitz, F. Vitamin-*A* and Sehpurpur: Ein Beitrag zur Kenntnis der Hemeralopie. *Med. Klin.*, 1933, **34**, 1148.

10. Hecht, S. Dark adaptation of the human eye. *J. Gen. Physiol.*, 1920, **2**, 499-517.

11. ———. The kinetics of dark-adaptation. *J. Gen. Physiol.*, 1927, **10**, 781-809.

12. ———. Vision: II. The nature of the photoreceptor process. In *A Handbook of General Experimental Psychology*, ed. by C. Murchison. Worcester: Clark Univ. Press, 1934. (Pp. 704-828.)

13. ———. Rods, cones, and the chemical basis of vision. *Physiol. Rev.*, 1937, **17**, 239-290.

14. Hecht, S., & Mandelbaum, J. Rod-cone dark-adaptation and vitamin-*A*. *Science*, 1938, **88**, 219.

15. Holm, E. Demonstration of hemeralopia in rats nourished on food devoid of fat-soluble-*A*-vitamin. *Amer. J. Physiol.*, 1925, **73**, 79-84.

16. Isaacs, B. L., Jung, F. T., & Ivy, A. C. Vitamin-*A* deficiency and dark-adaptation. *J. A. M. A.*, 1938, **111**, 777-780.

17. Jeans, P. C., & Zentmire, Z. A clinical method for the determination of moderate degrees of vitamin-*A* deficiency. *J. A. M. A.*, 1934, **102**, 892.

18. Jeghers, H. The degree and prevalence of vitamin-*A* deficiency in adults with a note on its experimental production in human beings. *J. A. M. A.*, 1937, **109**, 756.

19. Lashley, K. S. The mechanism of vision: V. The structure and image-forming power of the rat's eye. *J. Comp. Psychol.*, 1932, **13**, 173-200.

20. ———. The mechanism of vision: XVI. The functioning of small remnants of the visual cortex. *J. Comp. Neurol.*, 1939, **70**, 45-67.

21. Menner, E. Untersuchungen über die Retina mit besonderer Berücksichtigung der aüsseren Körnerschicht: Ein Beitrag zur Duplizitätstheorie. *Z. vergl. Physiol.*, 1928, **8**, 761-826.

22. Morgan, C. T. The dark-adaptation curve of normal rats and rats deprived of vitamin-A. *Psychol. Bull.*, 1938, **35**, 718.

23. ———. Studies in vision: I. A new technique for the study of visual discrimination in the rat. *J. Comp. Psychol.*, 1939, **28**, 73-79.

24. Munn, N. L. An investigation of color vision in the hooded rat. *J. Genet. Psychol.*, 1932, **40**, 351-362.

25. ———. Further evidence concerning color blindness in rats. *J. Genet. Psychol.*, 1934, **45**, 255.

26. Munn, N. L., & Collins, J. Discrimination of red by white rats. *J. Genet. Psychol.*, 1936, **48**, 72-87.

27. MUENZINGER, K. F., & REYNOLDS, H. E. Color vision in white rats: Sensitivity to red. *J. Genet. Psychol.*, 1936, **48**, 58-71.

28. PALMER, C. E. The dark-adaptation test for vitamin-*A* deficiency. *Amer. J. Publ. Hlth.*, 1938, **28**, 309-315.

29. SUGITA, Y. Studien über die physiologische und pathologische Verteilung der lipoid Substanze im Auge, speziell in der Netzhaut. *Arch. f. Ophth.*, 1925, **115**, 260-285.

30. TANSLEY, K. The regeneration of visual purple; its relation to dark-adaptation and night-blindness. *J. Physiol.*, 1931, **71**, 442-458.

31. WALD, G. Vitamin-*A* in the retina. *Nature*, 1933, **132**, 316.

32. WALD, G., & CLARK, A. Sensory adaptation and chemistry of the retinal rods. *J. Gen. Physiol.*, 1936, **21**, 93-105.

33. WALD, G., JEGHERS, H., & ARMINIO, J. An experiment in human dietary night-blindness. *Amer. J. Physiol.*, 1938, **123**, 732-746.

34. WALLS, G. L. The visual cells of the white rat. *J. Comp. Psychol.*, 1934, **18**, 363-366.

35. WALTON, W. Color vision and color preference in the albino rat. *J. Comp. Psychol.*, 1933, **15**, 373-394.

36. WATSON, J. B., & WATSON, M. I. A study of the responses of rodents to monochromatic light. *J. Anim. Behav.*, 1913, **3**, 1-14.

Laboratory of Physiological Psychology
Harvard University
Cambridge, Massachusetts

The Journal of Genetic Psychology, 1942, **60**, 121-136.

COMPARISON OF COPULATORY BEHAVIOR OF MALE RATS RAISED IN ISOLATION, COHABITATION, AND SEGREGATION*[1]

Laboratory of Experimental Biology, American Museum of Natural History

FRANK A. BEACH

A. INTRODUCTION

1. *Results of Earlier Investigations*

Steinach (1936) stated that male rats isolated for several months suffer testicular degeneration and consequent loss of copulatory reactions to the sexually-receptive female. If such males are subsequently permitted prolonged contact with female rats, testicular function is restored and mating behavior reappears. Stone (1922, 1923) reported normal sex drive in male rats raised in isolation from weaning to approximately 75 days of age. A few of these males experienced temporary difficulty of orientation to the female during initial periods of copulation. Comparable difficulty is displayed by sexually-inexperienced male rabbits (Macirone and Walton, 1938).

Male rats aroused by contact with the estrous female tend to become indiscriminate in their sexual aggression and to mount any other rat regardless of its sex or condition of receptivity (Stone, 1922; Lashley, 1938). Male rats accustomed to receiving receptive females in a certain cage tend to mount any other rat placed therein regardless of its behavior pattern (Jenkins, 1928; Beach, 1939). Comparable behavior is displayed by sexually-experienced male rabbits (Brooks, 1937; Edwards, 1940; Macirone and Walton, 1938).

When male animals are subjected to prolonged segregation homosexual mating attempts are frequent. This has been found true in the case of monkeys (Kempf, 1917), rabbits (Brown, 1937), guinea pigs (Avery, 1925; Louttit, 1927), and rats (Stone, 1924;

*Received in the Editorial Office on October 10, 1940.

[1]Supported by a grant from the Committee for Research on Problems of Sex, National Research Council, this experiment was begun in the Laboratory of Physiological Psychology at Harvard University and completed in the Laboratory of Experimental Biology at the American Museum of Natural History.

Jenkins, 1928; Beach, 1938). Male rats segregated for several months display a hyponormal heterosexual drive and a hypernormal homosexual drive as measured in the Columbia Obstruction Apparatus (Jenkins, 1928).

2. *Purpose of the Present Investigation*

The present work was undertaken in an attempt to answer four questions. (*a*) If male rats are raised in isolation from weaning to maturity do they show a weak sex drive? Steinach's results suggest a positive answer to this question. (*b*) When such rats first copulate with a receptive female is the mating pattern perfect? Stone's findings in this respect deal with males which were sexually but not physiologically mature. Further, the nature of Stone's experiments necessitated frequent observation periods during which the majority of his isolation cases were permitted contact with adult rats. (*c*) If male rats are raised with females how will their sexual behavior compare with that of other males deprived of heterosexual contact? Findings cited above indicate that experienced rats should show a strong sex drive and perfect copulatory pattern. (*d*) If male rats are raised in segregation from weaning to maturity how will their sexual behavior compare with that of isolated and of experienced males with respect to strength of drive, and accuracy of copulatory reactions? Earlier findings suggest that such males would display relatively weak sex drive; but there is no evidence regarding the heterosexual mating pattern displayed.

B. Experimental Methods

1. *Subjects and Colony Conditions*[2]

Fifty-five male rats 21 days old were divided by the split-litter technique into three experimental groups. Twenty-two males were raised in isolation from the day of weaning; and throughout the experiment these animals had no contact with other rats save during the copulation tests. Seventeen males were raised in segregation until they were 40 days old; and from this time until the beginning of the copulation tests 10 females were kept in the living cages with these experimental males. Females used for cohabitation with these

[2]During the latter part of the experiment assistance in colony maintenance was rendered by the personnel of the Works Progress Administration, Official Project No. 665-97-102.

males were sterilized by uterine ligation (Beach, 1938). Each week the used females were removed and fresh females were caged with the males. Pseudopregnant females were thus replaced by animals displaying normal estrous cycles. Sixteen males were segregated in one large cage from the day of weaning until the completion of the experiment.

All animals were raised under reversed light-dark conditions which brought the females into heat during the day and facilitated the conduction of sex tests at that time (Beach, 1938). Experimental animals were fed Purina Dog Chow, supplemented twice weekly with lettuce, and three times weekly with bread and milk.

2. *Copulation Tests*[3]

Sex tests were initiated when experimental males were 100 to 110 days of age. Three days before the first test the females were removed from the living cage of the cohabitation group, and thereafter these males had contact with females only during the observation periods.

Receptive females for use in copulation tests were selected on the basis of their reaction to non-experimental males of known sexual vigor. Experimental males were allowed a 15-minute pretest adaptation period in the observation cage. At the conclusion of this interval the receptive female was dropped into the cage, and the test lasted 15 minutes from the time of the first copulation. If no copulation occurred the test lasted 15 minutes from the introduction of the female. During the observation period the following data were kymographically recorded: (*a*) activation time,[4] (*b*) copulations, (*c*) attempts,[4] (*d*) vaginal plugs. Atypical methods of copulation were noted, and when no copulation occurred the male's responsiveness to the receptive female was rated on the following 3-point scale (Table A).

---

[3]The technique and apparatus employed in our observations of sexual activity in the male rat have been described and illustrated in an earlier paper (Beach, 1940).

[4]The number of seconds elapsing between the introduction of the female and the male's initial copulation is called the *activation time*. In an "attempt" the male mounts and palpates and may deliver pelvic thrusts. The general weakness of the response and the absence of the backward lunge associated with dismounting indicates that intromission was not achieved.

TABLE A

Rating	Behavior of the male
0	Pays no attention to female at any time.
1	Investigates female briefly once or twice during the test.
2	Exhibits mild sustained interest in the female, following her about the cage. Shows no evidence of excitement.
3	Persistent investigation of the female is accompanied by, or results in, definite excitement. Male jumps about the cage, climbs over the female frequently, pushes her against the sides of the cage, seems about to copulate. Male may mount the female but does not palpate or execute pelvic thrusts.

Occasionally a male attacked the receptive female and was given a rating of "fight." Mild biting of the female's ear, or pulling her about the cage with the teeth was not scored as fighting.

All experimental males were given seven copulation tests with a 3-day intertest rest. Twenty-one males which failed to copulate during the first seven tests were given 20 additional tests.

C. Experimental Results

1. *Proportion of Copulators in Each Group*

Table 1 indicates that during the first seven tests the highest percentage of copulators was found in the group raised in isolation. The group raised with females contained the next highest percentage of copulators; and the segregation group the lowest percentage.

TABLE 1

Per Cent of Each Group Copulating During the First Seven Tests

Group	*N*	Copulators	Per cent of group copulating
Isolation	22	15	69
Cohabitation	17	9	53
Segregation	16	4	25

Table 2 shows that 20 additional tests given to non-copulators from each group elicited copulatory behavior in the case of 100 per cent of the isolation animals, 80 per cent of the cohabitation males, and 45 per cent of the segregation cases. These data refer to only a part of the noncopulators in each group, for not all noncopulating cases could be given extra sex tests.

TABLE 2

PERFORMANCE OF NONCOPULATORS GIVEN 20 ADDITIONAL TESTS

Group	N given extra tests	N copulating	Per cent copulating
Isolation	3	3	100
Cohabitation	10	8	80
Segregation	11	5	45

Values presented in Table 2 suggest that if all animals had been given 27 instead of seven sex tests the actual number of copulators would have been higher, but intergroup relationships would not have been materially altered.

The percentage of each group copulating during each of the first six tests is shown in Table 3. These figures indicate that males

TABLE 3

PER CENT OF EACH GROUP COPULATING IN EACH OF THE FIRST SIX TESTS

Group	N	Per cent of group copulating Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Isolation	22	41	50	50	59	59	64
Cohabitation	17	24	35	24	35	24	29
Segregation	16	0	13	13	19	19	19

raised in isolation began copulating earlier in the test series than did animals raised under other conditions. Males raised with females copulated earlier than males raised in segregation. Since these are not cumulative scores the values do not coincide with data shown in Tables 1 and 2. The earlier tables deal with animals copulating at any time in the first seven tests; whereas the method of presentation employed in Table 3 permits fluctuations such as occur in the records of the cohabitation group, wherein on the third test the percentage of copulators is lower than it was on the preceding test.

Data included in Table 3 suggest that the intergroup differences shown in the preceding tables were established in the first sex test, and persisted throughout the series.

2. *Comparison of Copulators in each Group*

Table 4 shows that copulators raised in isolation displayed mating activity in a higher percentage of the first seven tests than did sexually

TABLE 4

COMPARISON OF COPULATORS IN THE FIRST SEVEN TESTS

	Isolation	Cohabitation	Segregation
Copulators	15	9	4
Tests Given	105	63	28
Tests Positive	85	38	16
Per Cent of Tests Positive	81	62	57

active males in the other two groups. Copulators raised in cohabitation copulated in a higher percentage of tests than did copulators raised with males.

Table 5 shows that the average number of precopulatory attempts

TABLE 5

PRECOPULATORY ATTEMPTS IN THE FIRST THREE POSITIVE TESTS FOR COPULATORS IN EACH GROUP

Group	Copulators	Positive Tests Precopulatory Attempts 1	2	3
Isolation	15	2.1*	0.8	0.5
Cohabitation	9	0.3	0.7	0.7
Segregation	4	0.2	1.8	1.2

*Elimination of the score of one male who attempted 22 times prior to the initial copulation reduces this average to 0.8.

in the first positive test was roughly equal for copulators in all three groups with the exception of one isolation animal which attempted 22 times before achieving copulation. In the second and third positive tests isolation and cohabitation copulators displayed nearly the same number of precopulatory attempts, while copulators raised in segregation attempted somewhat more frequently.

Table 6 indicates that copulators raised in isolation or in co-

TABLE 6

AVERAGE COPULATIONS PER POSITIVE TEST FOR COPULATORS IN EACH GROUP

Group	Copulators	Average copulations Positive tests First	Second
Isolation	14	9.4	14.5
Cohabitation	9	8.0	11.6
Segregation	4	10.0	10.0

habitation tended to deliver more copulations in the second positive test than in the first. Copulators raised in segregation showed no such increase in average number of copulations in successive positive tests.

In three successive positive tests 13 copulators raised in isolation delivered a total of 121, 181, and 186 copulations respectively. This finding, in harmony with data shown in Table 6, suggests that the number of copulations tends to increase in subsequent tests; but the curve levels off rapidly, and after the second or third test there is no significant increase.

3. *Responsiveness Ratings for each Group*

Table 7 and Figure 1 show that in negative tests males raised in

TABLE 7

RESPONSIVENESS OF ANIMALS IN EACH GROUP IN THE FIRST SEVEN TESTS

Group	Cases rated	Total ratings given	Per cent of ratings in each category 0	1	2	3	Fight
Isolation	12	54	7	9	11	56	17
Cohabitation	14	57	66	11	13	10	0
Segregation	15	70	74	1	7	7	1

isolation were more responsive to the receptive female than were males raised under other conditions. Males raised with females were slightly more responsive than males raised in segregation. During tests in which no copulation occurred isolation males received a low percentage of 0, 1, and 2 ratings and a high percentage of 3 ratings. Males raised in cohabitation and segregation received a high percentage of 0 ratings.

Eight males raised in isolation attacked the female. Four of these cases fought with the female during the first test; three others fought during the second test; and one isolation male attacked the receptive female in both the first and second tests. None of the males fought after the second test. One male raised in segregation attacked the female in the first test. Cohabitation males never fought with the female.

4. *Body Weight for each Group*

Table 8 indicates that males raised in isolation were heavier at the time of the first sex test than were animals in the other two

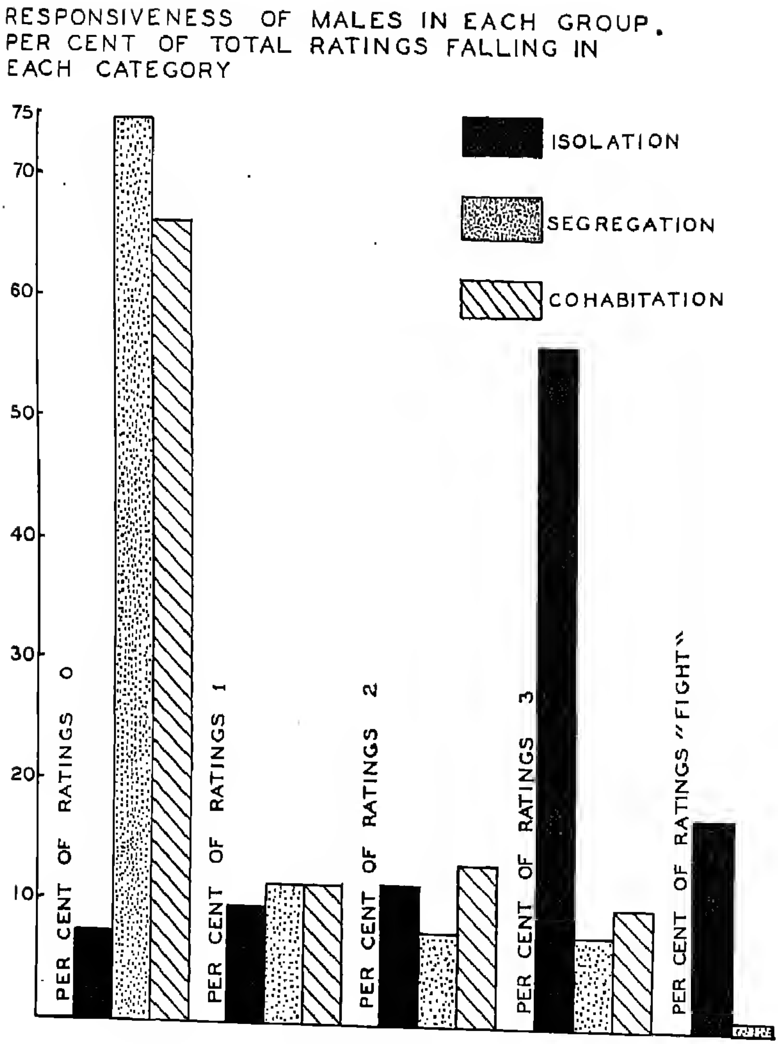

FIGURE 1

Comparison of Copulatory Behavior of Male Rats Raised in Isolation, Cohabitation, and Segregation

TABLE 8

AVERAGE WEIGHT OF ALL ANIMALS IN EACH GROUP AT THE TIME OF THE FIRST TEST (100-110 DAYS OLD) COMPARED TO NORMS PUBLISHED BY DONALDSEN (1924)

Group	*N*	Average	Grams weight Highest	Lowest
Isolation	22	326	380	260
Cohabitation	17	264	350	190
Segregation	16	251	335	160
Donaldsen	50	223	284	146

groups. Males raised in cohabitation were heavier than segregated males. All three groups show considerable overlap. Donaldsen (1924) found the average weight of stock albino males at 120 days of age to be 223 grams, which is much lighter than the lightest of our experimental groups. Donaldsen's figures are based upon animals of Wistar stock, from which our colony was originally derived. It appears therefore justifiable to draw the comparison between present data and those of Donaldsen.

The intergroup differences in body weight cannot be referred to a genetic difference for the groups were divided by the split-litter technique. Diet was constant for all groups. It follows that weight differences were the result of varying conditions under which the groups were raised.[5]

Table 9 reveals that copulators tended to be heavier than noncopulators within each group.

TABLE 9

AVERAGE WEIGHT OF COPULATORS AND NONCOPULATORS IN EACH GROUP DURING THE FIRST SEVEN TESTS

Group	Copulators *N*	Copulators Grams weight	Noncopulators *N*	Noncopulators Grams weight
Isolation	15	349	7	329
Cohabitation	9	316	8	247
Segregation	4	293	11	268

[5]In this connection it is of interest to note that Donaldsen obtained results directly opposed to our own. "The foregoing data," he writes, "represent the usual results under moderately favorable conditions. . . . Isolated animals grow less well than those kept together" (1924, p. 179).

D. Discussion

1. *Initial Sex Behavior of Males Raised in Isolation*

Craig observed that male doves raised in isolation often mounted the female in atypical fashion and displayed incomplete courtship patterns in initial sex tests.

> The sexual reaction of the male (dove) is, in its motor aspect, very completely and definitely provided in the innate nervous organization. When a dove performs an instinctive act for the first time, it generally shows some surprise, bewilderment, or even fear; and the first performance is in a mechanical, reflex style, whereas the same act after much experience is performed with ease, skill, and intelligent adaptation (1914, p. 132-133).

Eight sexually-inexperienced male rats observed by Stone executed their first copulation in a manner typical of an experienced copulator. Six other males mounted the female but did not palpate or exhibit pelvic thrusts. Fourteen males mounted and palpated without pelvic movement. Two young males mounted the female's head, showing palpation and pelvic thrusts. These observations lead Stone to conclude that, "in general the initial copulatory act of the young male resembles that of the experienced adult so closely that differences are not easily discernible" (1922, p. 123); and that his results substantiated "the generally accepted view that the copulatory response is the action of an hereditary mechanism" (ibid p. 147).

Stone added, however, that, "the copulatory pattern is a compound of both native and modified native or acquired responses" (ibid p. 148). Relative to the incomplete and disoriented initial mating responses of some animals Stone advanced the following explanation:

> The slight irregularities . . . may be variously explained. In some instances they are probably the result of physiological immaturity of some of the structures involved; in others, due to incomplete integration of the elementary activities of the copulatory unit or to incomplete dissolution of these elements from other distinctly foreign units with which they have functioned previously; or, again, they may arise from the failure of neural arcs of the copulatory act to gain priority over other arcs aroused by stimuli acting simultaneously with the stimulus adequate to evoke the copulatory act (1922, p. 147).

Twelve of the 15 copulators raised in isolation in the present experiment copulated normally during the first contact with the receptive female. Activation times and precopulatory attempts scored by these animals were comparable to those of copulators in other groups. Three additional males in this group mounted the female's head during the first test, and one animal attempted 22 times before achieving a complete copulation.[6] The first of Stone's explanations is not applicable in the present case for all males were physiologically mature. Since all copulators raised in segregation displayed only the normal mating pattern, the atypical behavior of three isolation cases cannot be referred solely to lack of heterosexual experience.

Head-mounting on the part of the isolation males appeared to be a product of two factors: (*a*) extreme excitement of the male, and (*b*) temporary inaccessibility of the receptive female. Usually the estrous female darts away from the investigating male and then crouches. When the pursuing male reaches her (and the direction of his approach is dictated by the fact that he *is* pursuing) he mounts from the rear; and with the placement of his forepaws upon her sides the female exhibits lordosis. None of the males in the present experiment head-mounted a female which had just run away. Head mounts occurred when the female's back was against the cage wall, and when she was therefore inaccessible for normal copulatory coupling.[7]

Abortive coupling is not restricted to inexperienced copulators. Recently we have observed that experienced males aroused by several copulations with the estrous female will occasionally head-mount the female when she backs into a corner during a period of temporary resistance.

In general present data with regard to initial copulatory behavior of male rats raised in isolation confirm the findings of Stone, suggesting that the inexperienced male is capable of normal copulatory reactions upon his first contact with the receptive female.

---

[6]It is important to note that during the tests in which these irregularities were observed the same males executed numerous normal copulatory acts.

[7]More recent experimentation strongly indicates that the size of the observation cage is an important factor affecting the amount of abortive mounting on the part of an excited male confronted with an actively receptive female. If there is room for the female to keep moving away from the *male so that her sexual posturing does not bring her "face-to-face"* with him, atypical coupling is greatly reduced or entirely eliminated.

2. *Comparison of Sex Drive in the Three Groups*

Our results are at variance with Steinach's report of loss of sex drive in isolated male rats (1936). Present findings indicate that males raised in isolation are more easily excited by the female than are males raised with other rats. Ball has noted that "several stimuli which seem to make the animal excited raise the level of sexual excitability" (1937, p. 3).

We have defined "responsiveness" in terms of the vigor and persistence of the male's investigation of the incentive animal. Isolation animals received higher responsiveness scores than did animals raised in segregation or with females. Males raised in isolation were more vigorous and persistent in their investigaton of the female; a fact which may logically be attributed to the novelty of contact with any other rat. An estrous female intensifies her receptive behavior in response to active investigation by the male. Increase in the amount of sexual posturing on the part of the female reciprocally heightens the male's excitement and increases the probability of copulation.

The excitement experienced by isolation males during their first contacts with another rat is reflected in the number of fights recorded. This behavior did not arise as a result of the female's refusal to mate, but usually followed a period of intense heat behavior on the part of the female and increased excitement on the part of the male. The rapid rise in the level of excitement culminated in an attack instead of a copulatory pattern. Rapid adjustment is indicated by a complete cessation of fighting after the second test. Males which attacked the female in early tests invariably copulated later in the series.

It may be suggested that, due to the novelty of contact with another rat, males raised in isolation conducted more extensive and intensive investigations of the stimulus animal than did males raised in segregation or cohabitation. In response to this persistent investigation by isolation males, the female increased the frequency and vigor of her sexual posturing. This intensification and prolongation of the female's heat behavior in turn elicited copulatory reactions in a relatively high percentage of the males.

In an experiment by Jenkins (1928) male rats segregated for several months were compared with nonsegregated controls with respect to their willingness to cross an electric grid to get to a receptive female. Segregated males scored fewer approaches, contacts

and crossings of the grid than did nonsegregated rats. When a second male was used as the incentive animal males raised in segregation crossed the grid more frequently than did nonsegregated cases. To explain these differences Jenkins suggested that homosexual tendencies were strengthened during prolonged segregation, lowering the heterosexual drive. Subsequent periods of cohabitation with females increased the heterosexual drive in some cases.

Since the present study involved no attempt to test the reaction of experimental males to other males we cannot confirm or disprove Jenkins' hypothesis. It seems clear, however, that the probability of copulation with a receptive female is definitely lower among rats raised in segregation than among males raised with females or raised in isolation.

3. *Body Weight and Sex Behavior*

All of our males appeared to be in good health. No cases of respiratory or labyrinth disorders were noted. The average weights of all three groups were above the normal weight reported by Donaldsen (1924). By using the split-litter technique in choosing our three groups we attempted to equate hereditary factors. All animals received the same diet, and an extra supply of dog chow was constantly available in all cages.

Despite these precautions isolation males gained more weight than cohabitation or segregation cases. Since copulators were usually heavier than noncopulators in each experimental group it may appear that intergroup differences in sexual activity are merely reflections of differences in body weight. Although greater weight may have been partially responsible for the high copulatory performance of isolation cases, the following relationships indicate that body weight was not the only important factor controlling intergroup differences: (*a*) Table 9 shows that seven isolation cases with an average weight of 329 grams failed to copulate; and that nine cohabitation cases and four segregation cases averaging 316 and 293 grams respectively did copulate. (*b*) Six males raised in segregation and tested every fourth day for seven months never copulated although their average weight during the last two months was 346 grams. (*c*) Table 3 shows that in the first six tests the proportion of segregation cases copulating increased from 0 to 19 per cent, and in this 24-day period the average body weight of these animals increased by less than five grams per rat.

E. Summary

Fifty-five male rats 21 days old were divided by the split-litter technique into three groups of 22, 17, and 16 cases. Animals in the first group were isolated in individual cages throughout the experiment. Males in the second group were raised in one cage to 40 days of age, and from that point until the beginning of the copulation tests this group was kept in a large cage with 10 sexually active females. Males in the third group were raised in segregation throughout the experiment.

Copulation tests with receptive females were initiated when the males were 100 to 110 days old, and occurred on every fourth day thereafter. Fifty-five males were given seven copulation tests; and 24 cases which failed to copulate during the first seven tests were given an additional 20 tests. Records made during each 15-minute copulation test included notes of activation time, number of copulations, attempts, and plugs. If no copulation occurred the male's responsiveness to the receptive female was rated on a 3-point scale.

The following results were obtained: (*a*) The proportion of copulators was highest among males raised in isolation, second among males raised with females, and lowest among males raised in segregation. (*b*) Copulating males raised in isolation tended to display mating reactions in a higher proportion of tests than did copulators raised under other conditions. Copulators raised with females achieved coition in a higher proportion of tests than did copulators raised in segregation. (*c*) The majority of males raised in isolation copulated perfectly upon their first contact with a receptive female; but in three cases abortive mounting occurred when the female was in an inaccessible position. (*d*) Males raised in segregation copulated perfectly when confronted with a receptive female. (*e*) During those tests in which no copulation occurred males raised in isolation tended to be more responsive to the female than did males raised with females. Males raised in cohabitation were more responsive than males raised in segregation. (*f*) Some males raised in isolation showed a tendency to attack the female during the first or second test. Such behavior rarely occurred in the case of males raised with other rats of either sex. (*g*) Males raised in isolation began copulating earlier in the test series than did nonisolated cases. Males raised with females copulated earlier in the series than males raised in segregation. (*h*) In general males raised in isolation weighed

more than males raised with females; and the latter tended to weigh more than males raised in segregation. (*i*) Copulators in each group tended to be heavier than noncopulators raised under similar conditions.

It is suggested that the high incidence of copulators in the isolation group is a product of two factors: (*a*) greater excitability resulting from the novelty of contact with a second animal; and (*b*) greater weight. The first factor involves prolonged investigation of the female by the male,—the resultant intensification of the female's estrous behavior,—reciprocal heightening of the male's state of excitement, and its eventual release in copulatory reactions.

The relative sexual inactivity of males raised in segregation may be partially explained on the basis of an increase in homosexual tendencies suggested by Jenkins (1928).

References

1. Avery, G. T. Notes on reproduction in guinea pigs. *J. Comp. Psychol.*, 1925, **5**, 373-396.
2. Ball, J. A test for measuring sexual excitability in the female rat. *Comp. Psychol. Monog.*, 1937, **14**, 1-37.
3. Beach, F. A. Sex reversals in the mating pattern of the rat. *J. Genet. Psychol.*, 1938, **53**, 329-334.
4. ———. The neural basis of innate behavior: III. Comparison of learning and instinctive behavior in the rat. *J. Comp. Psychol.*, 1939, **28**, 225-262.
5. Brooks, C. Mc. The rôle of the cerebral cortex and of various sense organs in the excitation and execution of mating activity in the rabbit. *Amer. J. Physiol.*, 1937, **120**, 544-553.
6. Brown, R. H. Stability of conditioning and sexual dominance in the rabbit. *Science*, 1937, **86**, 520.
7. Craig, W. Male doves reared in isolation. *J. Anim. Behav.*, 1914, **4**, 121-133.
8. Donaldsen, H. W. The Rat. *Mem. Wist. Instit.*, 1924, No. 6, 1-469.
9. Edwards, J. The effect of unilateral castration on spermatogenesis. *Proc. Roy. Soc.*, 1940, **128**, 407-421.
10. Jenkins, M. The effect of segregation on the sex behavior of the white rat as measured by the obstruction method. *Genet. Psychol. Monog.*, 1928, **3**, 457-471.
11. Kempf, E. J. The social and sexual behavior of the infra human primates with some comparable facts in human behavior. *Psychoanal. Rev.*, 1917, **4**, 127-154.

12. Lashley, K. S. Experimental analysis of instinctive behavior. *Psychol. Rev.*, 1938, **45**, 445-471.

13. Louttit, C. M. Reproductive behavior of the guinea pig: I. The normal mating behavior. *J. Comp. Psychol.*, 1927, **7**, 247-265.

14. Macirone, C., & Walton, A. Fecundity of male rabbits as determined by "dummy matings." *J. Agri. Sci.*, 1938, **28**, 122-134.

15. Steinach, E. Zur Gesichte des mannlichen Sexualhormon und seiner Wirking beim Saugetiere und beim Menschen. *Wein. Klin. Wschr.*, 1936, **49**, 161-172.

16. Stone, C. P. The congenital sexual behavior of the young male albino rat. *J. Comp. Psychol.*, 1922, **2**, 95-153.

17. ———. Further study of sensory functions in the activation of sexual behavior in the young male albino rat. *J. Comp. Psychol.*, 1923, **3**, 469-473.

18. ———. A note on "feminine" behavior in adult male rats. *Amer. J. Physiol.*, 1924, **68**, 407-424.

Laboratory of Experimental Biology
American Museum of Natural History
Central Park West & 79th Street
New York City

The Journal of Genetic Psychology, 1942, **60**, 137-142.

EXECUTION OF THE COMPLETE MASCULINE COPULATORY PATTERN BY SEXUALLY RECEPTIVE FEMALE RATS*

Laboratory of Experimental Biology, American Museum of Natural History

FRANK A. BEACH

Female rats in estrous occasionally respond to other sexually-receptive females with the execution of portions of the masculine copulatory pattern (Long and Evans, 1922; Hemmingsen, 1933; Stone, 1938; Ball, 1940). One published report describes sexually-receptive female rats which directed masculine mating reactions toward sluggish males that had failed to respond to the female's sexual posturings (Beach, 1938).

The complete sex pattern of the male rat first described by Stone (1922) and verified by other workers (Ball, 1939; Anderson, 1936; Beach, 1940) consists of five elements. (*a*) The male mounts the female from the rear, clasping her sides between his forepaws. (*b*) Moving the forepaws inward, backward and upward, and then releasing the pressure, the male palpates the female's sides with extreme rapidity. (*c*) Moving the hind quarters in and out in a piston-like fashion the male delivers a number of vigorous pelvic thrusts which presumably direct the penis into the vagina. (*d*) After the final pelvic thrust the male's forepaws release the female and he dismounts with a forceful backward lunge. The backward lunge often carries the male 6 to 10 inches from the female, and an extremely vigorous copulator may fall over backward or carom off the cage wall. (*e*) Coming to rest in a sitting position the male licks the penis and scrotum for several seconds, often manipulating the external genitals with the forepaws.

Observers who have described in detail the masculine sexual behavior of normal females have emphasized the absence of the fourth and fifth elements characteristic of the copulating male's mating pattern. No account has come to our attention in which normal females are described as displaying the backward lunge or the post-copulatory licking of the genitals.

*Received in the Editorial Office on October 10, 1940.

To induce masculine sexual behavior in female rats Ball (1940) injected 10 adult females with testosterone propionate and studied their reactions to normal females in heat. Females receiving the male hormone showed masculine mating reactions more frequently than did untreated females. Further, injected females performed the complete male pattern including the backward lunge and postcopulatory genital cleaning.

The purpose of the present paper is to report the execution of the complete masculine pattern by untreated female rats. This behavior was noted in the course of a current experiment dealing with the copulatory behavior of male rats.[1] At the beginning of each day's testing three to six sexually-receptive females, to be used as incentive animals with experimental males, are selected from stock. A female's receptivity is tested by placing her with a nonexperimental male of known sexual vigor. Each female may be used in several tests with different experimental males during the day's work. Copulation with a vigorous male frequently appears to increase a female's excitability so that when she is returned to the resting cage containing other estrous females she resorts to the masculine mating pattern.

Recently we were testing a series of sexually active males, using three highly receptive females as incentive animals. These females were multiparous rats 10 to 13 months old. During the intertest rests we noted that two females were alternately mounting the third female. One of the aggressive females was seen to display the backward lunge and postcopulatory genital cleaning; and when this occurred an immediate attempt was made to induce repetition of the behavior and to record it in motion pictures.

The aggressive female which had executed the complete male pattern was placed in the regular observation cage and allowed a few minutes' rest. The submissive female was then placed in the observation cage, and in the next 30 minutes the aggressive female mounted her cage mate nine times. Each time that mounting occurred the bright lights were turned on for movie work; but this stimulus disrupted the behavior in question and the mounting female immediately released her partner to retire to a far corner of the cage. For the next 10 minutes the camera was put aside and

[1]Supported by a grant from the Committee for Research in Problems of Sex, National Research Council.

brilliant lights were not used. In this period eight partial mountings and three complete acts were observed.

The completeness of the act was unquestionable, and the backward lunge was quite pronounced, carrying the mounting female several inches from the animal she had just released. Each time the backward lunge occurred the copulating female assumed a sitting posture and assiduously licked the genital region in a manner characteristic of the copulating male. In a final attempt to obtain a permanent record we were able to film one complete act in which the backward lunge, though present, was somewhat weaker than usual.

At the conclusion of these observations the submissive female was returned to the resting cage, and immediately the second aggressive female mounted and displayed the complete pattern twice. Although we watched the two animals for a half hour this behavior did not recur. Mounting and palpation were frequent, but the backward lunge and postcopulatory genital cleaning did not reappear.

In an earlier study estrous females exhibiting the male sexual pattern were found to possess macroscopically normal ovaries (Beach, 1938). That estrogenic activity in the ovaries of the females herein described was normal is indicated by the animals' marked sexual receptivity when confronted with a male. A partial explanation for the behavior observed appears to lie in a suggestion previously advanced:

> . . . the specificity of the mating patterns for the two sexes, although probably inherited, is not rigidly dictated by the innately organized substratum. Although there may be a strong preference for the normal copulatory response it is obvious that in a few individuals at least, there exists the innate organization essential to the mediation of the mating pattern of either sex (Beach, 1938, p. 334).

In describing the effects of testosterone upon female rats Ball wrote as follows:

> It is concluded that the male copulatory pattern in more or less rudimentary form is part of the equipment of the normal female rat. The threshold of this behavior is very high normally, but it can be lowered by testosterone administration (1940, p. 164).

Although Ball drew no such conclusion, one might have added the conjecture that in addition to lowering the threshold for the rudi-

mentary male pattern in the female, the testosterone contributed an exclusively masculine factor embodied in the two final elements of the copulatory act. Present results indicate that this is not the case. The complete male pattern is present in some untreated females; and occasionally may be evoked in response to a second receptive female.

In this connection it is significant to note that Noble and Wurm (1940) have recorded one instance of a female black-crowned night heron (with histologically normal ovaries) which exhibited elements of the mating pattern normally restricted to the behavior of the courting male.

The reaction of male and female herons to injections of testosterone propionate lead these authors to the following conclusions:

> Differences in the sexual behavior of the adults seem regulated only by differences in the amounts of male hormone normally found in these birds. Estrogens alone fail to stimulate any breeding behavior in either sex (1940, pp. 849-850).

Since castrate birds failed to develop any secondary sex characters it was concluded that the androgens responsible for both male and female courtship are produced in the ovary.

The production of androgenic substances by the mammalian ovary has been demonstrated in the case of the mouse (Hill, 1937, *a*, *b*), pig (Parkes, 1937), and rat (Deansley, 1938). An alternative source of androgens in the female may be the suprarenal cortex. Allen and Vespignani (1938) found active testicular epithelium in the connective tissue surrounding the adrenal. Virilism in the human female is often related to excessive androgenic material revealed by urine assay, and the masculine characters accompanying such a condition may be reduced or eliminated by unilateral adrenalectomy (Broster, Allen, Vines, Patterson, Greenwood, Marrian, and Butler, 1938). Derivatives of the adrenal cortical hormone exert an androgenic effect (Mason and Myers, 1936); and administration of adrenal tissue effects masculinization of the female guinea pig (Hodler, 1937).

Female rats herein described may have reacted to excessive androgenic activity on the part of the ovaries or adrenals.

The literature includes two reports of virile male rats which displayed portions of the female copulatory pattern when mounted by other sexually-active males (Stone, 1924; Beach, 1938). These

animals may have reacted under the influence of excessive amounts of estrogens secreted by the testes. Urine assays reveal that the human testis produces estrogens (Witschi, 1939); and the urine of male homosexuals has been found to contain a higher proportion of estrogenic material than the urine of normal men (Glass, Duel, and Wright, 1940).

Temporary sex reversals in the mating pattern of male and female rats cannot be fully explained on the basis of estrogens in the male and androgens in the female. Such hypotheses, even if they could be fully substantiated in each case with endocrinological evidence, fail to take account of the neural basis of the behavior involved. When an adult male rat, raised in isolation, executes perfect copulatory behavior upon his first contact with a receptive female, we are forced to assume the existence of an inherited neural organization mediating the complex concatenation of reflexes. The same is true of sexually receptive behavior displayed by the inexperienced female during the first estrous.

Findings cited in the present report suggest that some individuals may inherit or acquire the neural organization mediating the mating pattern of the opposite sex. The appearance of this pattern is apparently dependent upon its threshold, and the stimulation offered by the behavior of the incentive animal. No complete explanation for the observed facts is yet available, but they appear to be important to a complete understanding of sexual behavior, and must be taken into account in the interpretation of "sex reversals" in mammals subjected to hormone administration.

References

1. Allen, E., & Vespignani, P. M. Active testicular epithelium in the connective tissue surrounding a human suprarenal gland. *Anat. Rec.*, 1938, **72**, 293.
2. Anderson, E. Consistency of tests of copulatory frequency in the male albino rat. *J. Comp. Psychol.*, 1936, **21**, 447-459.
3. Ball, J. Male and female mating behavior in prepubertally castrated male rats receiving estrogens. *J. Comp. Psychol.*, 1939, **28**, 273-283.
4. ———. The effect of testosterone on the sex behavior of female rats. *J. Comp. Psychol.*, 1940, **29**, 151-165.
5. Beach, F. A. Sex reversals in the mating pattern of the rat. *J. Genet. Psychol.*, 1938, **53**, 329-334.
6. ———. Effects of cortical lesions upon the copulatory behavior of male rats. *J. Comp. Psychol.*, 1940, **29**, 193-245.

7. BROSTER, L. R., ALLEN, C., VINES, H. W. C., PATTERSON, J., GREENWOOD, A. W., MARRIAN, G. F., & BUTLER, G. C. The adrenal cortex and intersexuality. London: Chapman & Hall, 1938.

8. DEANSLEY, R. The androgenic activity of ovarian grafts in castrated male rats. *Proc. Roy. Soc. London*, 1938, **126**, 122.

9. GLASS, S. J., DUEL, H. J., & WRIGHT, C. A. Sex studies in male homosexuality. *Endocrinology*, 1940, **26**, 590-594.

10. HEMMINGSEN, A. Studies on the oestrous-producing hormone (Oestrin). *Skand. Arch. f. Physiol.*, 1933, **66**, 97-250.

11. HILL, R. T. Ovaries secrete male hormone. I. Restoration of the castrate type of seminal vesicle and prostate glands to normal by grafts of ovaries in mice. *Endocrinology*, 1937*a*, **21**, 495.

12. ———. Ovaries secrete male hormone. III. Temperature control of male hormone output by grafted ovaries. *Endocrinology*, 1937*b*, **21**, 633.

13. HODLER, D. Surrénales et masculinisation. *Arch. Anat., Hist. & Embr.*, 1937, **24**, 1.

14. LONG, J. A., & EVANS, H. M. The oestrous cycle in the rat and its associated phenomena. *Mem. Univ. Calif.*, 1922, **6**.

15. MYERS, C. S., & KENDALL, E. C. Chemical studies of the suprarenal cortex. *J. Biol. Chem.*, 1936, **116**, 267.

16. NOBLE, G. K., & WURM, M. The effect of testosterone propionate on the black-crowned night heron. *Endocrinology*, 1940, **26**, 837-850.

17. PARKES, A. C. Androgenic activity of ovarian extracts. *Nature*, **139**, 965.

18. STONE, C. P. The congenital sexual behavior of the young male albino rat. *J. Comp. Psychol.*, 1922, **2**, 95-153.

19. ———. A note on "feminine" behavior in adult male rats. *Am. J. Physiol.*, 1924, **68**, 39-41.

20. ———. Personal communication to the author (1938).

21. ———. Sex drive. In *Sex and Internal Secretions*. (Eds. Allen, Danforth, & Doisy.) Baltimore: Williams & Wilkins, 1939.

22. WITSCHI, E. Modification of development of sex in lower vertebrates and in mammals. In *Sex and Internal Secretions*. (Eds. Allen, Danforth, & Doisy.) Baltimore: Williams and Wilkins, 1939.

Laboratory of Experimental Biology
American Museum of Natural History
Central Park West & 79th Street
New York City

The Journal of Genetic Psychology, 1942, **60**, 143-151.

SOME PERSONALITY TRAITS OF HARD OF HEARING CHILDREN*

Teachers College, Columbia University

R. Pintner[1]

Partial loss of hearing among school children is known to make it somewhat more difficult for the hard of hearing child to keep up with his regular classmates in school. A previous study by the present writer (2) suggests that hard of hearing children do not feel so well-adjusted in school. The present study attempts to go further in the comparison of hard of hearing and normal hearing children with reference to three personality traits.

The Subjects

During a three-year period from January, 1934, to June, 1938, the test, *Aspects of Personality* (3), was administered to 1,171 hard of hearing children and 1,208 normal hearing children in the public schools of New York City.

The hard of hearing children were selected from the records maintained by the New York League for the Hard of Hearing. The hearing loss of these children was measured by the *4A* audiometer, and the *2A* audiometer, and finally checked by individual otological examinations. Normal hearing children from the same grades were selected to be used as controls. Only those most recently tested for hearing were chosen for this study. The normal or control group was selected at random from the same class as the hard of hearing.

The data thus gathered have been considered in four samples. Sample I consists of a group of normal and hard of hearing children in Grades V and VI examined before July, 1936. The hard of hearing were taken from 26 schools in the Bronx, Brooklyn, Manhattan, and Queens. The total number of children in *Sample I* is:

*Received in the Editorial Office on October 12, 1940.

[1]The writer wishes to acknowledge the general help received in this study by Project No. 655-97-3-6 of the Works Progress Administration, and in particular he wishes to mention the specific aid given by Miss Beatrice G. Koot who, as a worker on the project, was responsible for much of the statistical work.

	Hard of hearing	Normal
Boys	289	268
Girls	315	212

Sample II: The children in this sample were examined from October, 1936, to June, 1937. Children in Grades V and VI were selected from 35 schools in the Bronx, Brooklyn, and Manhattan. The number in this sample is:

	Hard of hearing	Normal
Boys	84	152
Girls	74	82

Sample III: The children in this sample were studied from September, 1937, to June, 1938. Children in Grades V and VI were selected from 21 schools in the Bronx, Brooklyn, and Manhattan. The total number in this sample is:

	Hard of hearing	Normal
Boys	116	120
Girls	73	95

Sample IV: For this sample the same schools were used as for Sample III. However, only seventh and eighth grade children were chosen. The total number in this sample is:

	Hard of hearing	Normal
Boys	93	117
Girls	127	162

A total of 2,379 children was examined of which 1,171 were hard of hearing and 1,208 were normal hearing. There were 582 hard of hearing boys and 589 hard of hearing girls. There were 657 normal hearing boys and 551 normal hearing girls. In all comparisons of normal and hard of hearing we have divided the hard of hearing into two sub-groups, namely those with less than 15 decibels loss in the better ear and those with 15 or more decibel loss. In the final comparisons we have further selected out a small group having 30 or more decibels loss in the better ear. This sub-group is included in the group having 15 or more decibels loss.

The Test

The *Aspects of Personality Test* was administered to all the children both normal and hard of hearing. This test depends on oral and printed directions. It measures three different phases of personality, Ascendance-Submission, Extroversion-Introversion, and Emotionality. The Ascendance-Submission part measures the child's tend-

ency to dominate or be dominated. A high score on this part indicates ascendance. The Extroversion-Introversion part attempts to differentiate between the extrovert and the introvert. A high score indicates extroversion. The Emotionality part measures the emotional adjustment of the child. A high score indicates that a child is well adjusted, a low score that he is not.

Ascendance-Submission

Table 1 presents the results for the four samples for ascendance-

TABLE 1
Ascendance-Submission: Means and Sigmas for Ascendance-Submission Scores

Sample	Boys N	Boys M	Boys σ	Girls N	Girls M	Girls σ
I. Normal hearing	268	18.7	4.8	212	16.4	4.7
Hard of hearing less than 15	176	18.7	4.1	153	16.2	4.7
All hard of hearing	289	18.9	4.2	315	16.4	4.7
Hard of hearing 15 or more	113	19.1	4.4	162	16.6	4.7
II. Normal hearing	152	17.3	4.5	82	14.3	4.6
Hard of hearing less than 15	55	18.0	4.3	53	14.8	3.5
All hard of hearing	84	18.0	4.3	74	15.2	3.6
Hard of hearing 15 or more	29	17.9	4.2	21	16.1	3.8
III. Normal hearing	120	18.5	4.3	95	16.2	5.1
Hard of hearing less than 15	75	17.9	4.4	30	16.4	5.0
All hard of hearing	116	17.8	4.2	73	16.5	5.0
Hard of hearing 15 or more	41	17.7	3.7	43	16.5	5.1
IV. Normal hearing	117	18.6	4.7	162	17.5	4.5
Hard of hearing less than 15	53	17.2	3.6	60	16.8	4.4
All hard of hearing	93	17.9	4.2	127	16.9	4.3
Hard of hearing 15 or more	40	18.8	4.7	67	17.0	4.2

submission. The differences between the normal and various hard of hearing groups are very small. There seems to be a suggestion that those with most hearing loss score highest. Girls, both normal hearing and hard of hearing, make lower scores, i.e., are more submissive, than boys.

TABLE 2
ASCENDANCE-SUBMISSIONS COMPARISON OF MEANS AND SIGMAS FOR NORMAL HARD OF HEARING

Sample	N	M	σ	$M_{diff.}$	t
Boys					
Normal hearing	657	18.31	4.67		
Hard of hearing less than 15	359	18.20	4.17	.11	.35
All hard of hearing	582	18.37	4.23	.06	.20
Hard of hearing 15 or more	223	18.64	4.32	.33	.85
Hard of hearing 30 or more	17	18.71	4.81	.40	.34
Girls					
Normal hearing	551	16.37	4.83		
Hard of hearing less than 15	296	16.09	4.55	.28	.83
Hard of hearing 15 or more	293	16.65	4.57	.28	.85
All hard of hearing	589	16.37	4.57	.00	.00
Hard of hearing 15 or more	293	16.65	4.57	.28	.85
Hard of hearing 30 or more	31	17.45	3.89	1.08	1.27

Table 2 presents the results for the four samples combined, again keeping the boys and girls separate. An examination of this table shows a slight tendency for the mean scores to increase as we go from the normal to the extreme hard of hearing groups. This tendency is, however, very slight and there are no significant differences.

Table 3 shows the correlations between scores for ascendance-

TABLE 3
CORRELATIONS BETWEEN ASCENDANCE-SUBMISSION AND HEARING LOSS

Sample		N	Correlation with better ear	Correlation with poorer ear
I.	Boys	289	.03	—.12
	Girls	315	.07	.04
II.	Boys	84	.07	—.09
	Girls	74	.16	.15
III.	Boys	116	—.05	—.01
	Girls	73	.05	.02
IV.	Boys	93	.15	—.04
	Girls	127	.01	—.01
	All boys	582	.05 SD_r .04	—.08 SD_r .04
	All girls	589	.08 SD_r .04	.05 SD_r .04

submission and decibel loss in both the poorer and the better ear. All of the correlations are practically zero. We may therefore conclude from a study of the mean scores for various groups and of

the correlations, that there is no relation between ascendance-submission and amount of hearing loss.

Extroversion-Introversion

Tables 4, 5, and 6 present the data for the extroversion-introversion

TABLE 4

Extroversion-Introversion: Means and Sigmas for Extroversion-Introversion Scores

Sample	Boys N	Boys M	Boys σ	Girls N	Girls M	Girls σ
I. Normal hearing	268	22.0	3.7	212	21.7	3.7
Hard of hearing less than 15	176	21.8	3.8	153	20.9	3.6
All hard of hearing	289	21.4	3.9	315	21.0	3.6
Hard of hearing 15 or more	113	21.0	3.9	162	21.1	3.6
II. Normal hearing	152	21.9	4.3	82	20.6	3.8
Hard of hearing less than 15	55	21.5	3.9	53	19.9	3.7
All hard of hearing	84	21.8	4.2	74	20.0	4.0
Hard of hearing 15 or more	29	22.2	4.6	21	20.1	4.7
III. Normal hearing	120	21.3	3.5	95	19.6	4.0
Hard of hearing less than 15	75	19.8	4.3	30	19.5	4.2
All hard of hearing	116	20.1	3.9	73	19.5	4.1
Hard of hearing 15 or more	41	20.7	3.1	43	19.6	4.0
IV. Normal hearing	117	21.6	4.1	162	20.5	4.5
Hard of hearing less than 15	53	20.9	4.7	60	21.4	3.8
All hard of hearing	93	21.0	4.6	127	20.6	4.2
Hard of hearing 15 or more	40	21.2	4.5	67	20.0	4.3

scores. So far as this trait is measured by a pencil and paper group test, we can detect no difference between hard of hearing and normal hearing children. Even those with comparatively great hearing losses (i.e., 30 or more decibels) do not show any tendency to withdraw within themselves and become introverted.

Emotionality

The results for emotional stability are shown in Tables 7, 8, 9. An examination of these tables shows little difference between the

TABLE 5

EXTROVERSION-INTROVERSION: COMPARISON OF NORMAL AND HARD OF HEARING BOYS AND GIRLS

Sample	N	M	σ	$M_{diff.}$	t
Boys					
Normal hearing	657	21.79	3.87		
Hard of hearing less than 15	359	21.19	4.13	.60	2.13
All hard of hearing	582	21.16	4.09	.63	2.51
Hard of hearing 15 or more	223	21.13	4.01	.66	2.20
Hard of hearing 30 or more	17	21.71	4.51	.08	.09
Girls					
Normal hearing	551	20.82	4.09		
Hard of hearing less than 15	296	20.68	3.78	.14	.47
All hard of hearing	589	20.61	3.87	.21	.74
Hard of hearing 15 or more	293	20.54	3.97	.28	.97
Hard of hearing 30 or more	31	19.36	3.69	1.46	1.94

TABLE 6

CORRELATIONS BETWEEN EXTROVERSION-INTROVERSION AND HEARING LOSS

Sample		N	Correlation with better ear	Correlation with poorer ear
I.	Boys	289	—.05	.02
	Girls	315	—.02	.02
II.	Boys	84	.13	.03
	Girls	74	.03	—.02
III.	Boys	116	.12	—.16
	Girls	73	—.12	—.10
IV.	Boys	93	—.004	—.08
	Girls	127	—.26	—.06
	All boys	582	.004 SD_r .04	—.03 SD_r .04
	All girls	589	—.08 SD_r .04	—.01 SD_r .04

normal hearing and hard of hearing groups in general. However, the two small groups having the greatest hearing loss of 30 decibels or more show the lowest averages. There would seem to be a suggestion here that when the hearing loss becomes fairly severe lack of emotional stability becomes measurable by means of our test.

INTELLIGENCE

Scores on the Pintner *Intelligence Test* were available for 1,089 hard of hearing cases out of the 1,171 hard of hearing cases tested by the *Aspects of Personality Test*. The Pintner *Intelligence Test*

TABLE 7

EMOTIONALITY: MEANS AND SIGMAS FOR EMOTIONALITY SCORES

Sample	Boys *N*	Boys *M*	Boys σ	Girls *N*	Girls *M*	Girls σ
I. Normal hearing	268	25.8	6.5	212	26.9	5.9
Hard of hearing less than 15	176	25.4	6.7	153	27.0	5.9
All hard of hearing	289	24.7	6.6	315	26.2	6.1
Hard of hearing 15 or more	113	23.7	6.3	162	25.4	6.1
II. Normal hearing	152	24.4	6.9	82	25.2	6.9
Hard of hearing less than 15	55	23.3	6.5	53	24.2	6.4
All hard of hearing	84	23.6	6.6	74	24.6	6.4
Hard of hearing 15 or more	29	24.1	6.8	21	25.6	6.4
III. Normal hearing	120	24.3	7.1	95	25.5	4.6
Hard of hearing less than 15	75	22.8	6.8	30	24.7	7.9
All hard of hearing	116	23.0	7.2	73	23.8	7.2
Hard of hearing 15 or more	41	23.3	7.5	43	23.2	6.7
IV. Normal hearing	117	26.5	6.2	162	26.7	5.5
Hard of hearing less than 15	53	26.3	6.9	60	27.6	5.5
All hard of hearing	93	25.6	6.6	127	27.0	5.7
Hard of hearing 15 or more	40	24.7	6.2	67	26.6	5.8

TABLE 8

EMOTIONALITY: COMPARISON OF NORMAL AND HARD OF HEARING BOYS AND GIRLS

Sample	*N*	*M*	σ	$M_{diff.}$	*t*
Boys					
Normal hearing	657	25.34	6.72		
Hard of hearing less than 15	359	24.66	6.84	.68	1.41
All hard of hearing	582	24.36	6.71	.98	2.07
Hard of hearing 15 or more	223	23.87	6.59	1.47	2.85
Hard of hearing 30 or more	17	22.24	6.82	3.10	1.82
Girls					
Normal hearing	551	26.35	5.97		
Hard of hearing less than 15	296	26.40	6.29	.05	.10
All hard of hearing	589	25.88	6.32	.47	1.21
Hard of hearing 15 or more	293	25.37	6.31	.98	2.28
Hard of hearing 30 or more	31	21.03	5.03	5.32	5.58

TABLE 9

CORRELATIONS BETWEEN EMOTIONALITY AND HEARING LOSS

Sample		N	Correlation with better ear	Correlation with poorer ear
I.	Boys	289	—.11	.05
	Girls	315	.18	—.13
II.	Boys	84	.02	.18
	Girls	74	—.04	.02
III.	Boys	116	.01	—.11
	Girls	73	—.10	—.04
IV.	Boys	93	—.20	—.20
	Girls	127	—.20	—.18
	All boys	582	—.07 SD_r .04	.01 SD_r .04
	All girls	589	—.14 SD_r .04	—.11 SD_r .04

[now called the Pintner *General Ability Test* (1)] is a group verbal intelligence test. Correlations between intelligence and the three traits for boys and girls separately and for the total 1,089 cases were calculated. All of these correlations are very low. The correlations for the total group, including both boys and girls, are as follows:

Intelligence and Ascendance-Submission	= +.06
Intelligence and Extroversion-Introversion	= +.06
Intelligence and Emotionality	= +.28

Correlations between decibel loss in the better ear and intelligence score were also calculated. For the total group this correlation is +.04.

COMPARISON WITH NORMS

In the manual for the *Aspects of Personality Test* percentile norms are given by the authors of the test. A comparison has been made of the mean scores of our groups with these norms. The Ascendance-Submission scores for our various groups are all somewhat above the fifty percentile of the norms. The Extroversion-Introversion scores are all close to the fifty percentile of the norm. The mean score on Emotionality is at the fifty percentile for the girls but slightly below this point for the boys. The only decided deviation from the norms is found on Emotionality for the group with hearing loss of 30 decibels or more. The mean scores for both boys and girls in this group fall at about the 25th percentile of the standardization

group. This would seem to indicate more emotional instability among hard of hearing children with extreme hearing losses.

Summary

A group personality test was given to 1,171 hard of hearing and 1,208 normal hearing children in Grades V to VIII inclusive. From many comparisons of various samples it would appear that there is fundamentally no difference between normal hearing and hard of hearing children in such traits as ascendance-submission and extroversion-introversion so far as these traits can be measured by a group test of the inventory type. With regard to emotional stability there would appear to be a tendency for the more extreme cases of hearing loss to score lower on this test. Such cases may on the average be slightly less emotionally stable than normal hearing children or than children whose loss of hearing does not exceed 30 decibels.

References

1. PINTNER, R. Pintner General Ability Test. New York: World Book, 1939.
2. ———. An adjustment test with normal and hard of hearing children. *J. Genet. Psychol.*, 1940, 56, 367-381.
3. PINTNER, R., *et al.* Aspects of Personality Test. New York: World Book, 1938.

Teachers College
Columbia University
New York City

The Journal of Genetic Psychology, 1942, 60, 153-165.

EXPERIMENTAL ANALYSIS OF THE VASELINE TECHNIQUE OF KUO FOR STUDYING BEHAVIORAL DEVELOPMENT IN CHICK EMBRYOS*[1]

Anatomical Laboratories, Northwestern University Medical School

R. FREDERICK BECKER

A. INTRODUCTION

A difference of opinion still exists among investigators as to the nature of early fetal behavior. One school of thought (1, 4, 8) describes reflex activity as individuating from a primary, integrated total pattern. Another group of workers (3, 22, 25) believe that simple, localized, movements are the first manifestations of behavior, and only later, along with further growth of the nervous system, become integrated into more complex patterns. The subject has been reviewed recently (20). In this as in some other fields of biological investigation, much that is controversial might be unified if sufficient attention were given to a few fundamental principles.

For example, the conditions under which an experiment is performed should be physiologically sound. Most experiments involving embryos and fetuses must be carried out in an environment which corresponds as closely as possible to that obtaining in ovo or in utero. Situations in the experimental environment which differ from the normal environment demand careful consideration when it comes to evaluating results.

Species differences also should be given careful consideration. It cannot be assumed that prenatal physiology and ecology are the same for all animals, rather such differences as exist need to be recognized in generalizing from the behavior of one species to another. However, it does seem logical to expect greater similarity in behavior, even during fetal life, between forms closely related phylogenetically than between species widely separated in this respect.

All factors which influence the developing organism, whether they be external or internal, are important considerations in an understanding of the development of behavior. It cannot be assumed that

*Received in the Editorial Office on October 14, 1940.

[1]This investigation was made possible by a grant-in-aid of research from the John and Mary R. Markle Foundation to Dr. W. F. Windle.

the external environment of the egg or the uterus alone determines the nature of the early behavioral reactions. True, it may play an important rôle at times in modifying, or even in initiating, the behavior of the organism. Nevertheless, it is equally as important to take into account the state of the developing nervous system within the organism, and the changes in the *"milieu interne"* which may stimulate, inhibit, or modify nervous and muscular function. The tendency usually has been to emphasize one of these factors to the exclusion of others.

The purpose of this study is to examine critically the experimental method used by Kuo (9-18) and its bearing upon embryonic behavior and the reflex controversy. The conditions under which Kuo's experiments were performed, as reported in 1932 (9, 10), were as follows.

The shell of the incubating chick's egg was opened in the region of the respiratory air space, thus exposing part of the inner-shell membrane which was coated over with a thin layer of liquid vaseline, to render it transparent. The eggs were then observed daily while they incubated upright on their small ends. The impression is given that the behavior of the same chick could be observed daily from the second, or earliest day of operation, to the twenty-first, or day of hatching. In all articles subsequent to 1932 (except one paper in collaboration with Carmichael in 1937), Kuo has referred to this original procedure (9, 10) as the standard for his experimental method. Behavior observed under such conditions he has regarded as normal.

The question arises as to the physiologic soundness of an experimental environment in which the chick's respiratory exchange is limited by coating a large portion of the shell membrane with vaseline. Development, even in early stages, must be forced to progress under conditions of anoxemia. Such an assumption tends to receive support from the early experiments of Baudrimont and St. Ange (2) and Dareste (5-7) who varnished various regions of the shells of incubating eggs. Moreover, in mammalian work it has been established that the blood-gas relationships of the fetus must be disturbed as little as possible if reliable behavioral observations are desired (22). Evidently, this very question of oxygen want once was considered by Kuo, for in a paper with Carmichael (18), it is noted that only a spot directly over the embryo instead of the entire shell membrane was vaselined in order not to hamper respiration.

However, the chicks in this particular study were used only temporarily for cinematographic purposes, and the number reaching the hatching stage was not stated. In fact, while the number of chicks used by Kuo in any experiment often ran in the hundreds, he has never recorded the number of chicks which hatched or the number which failed to do so. The present experiments were undertaken to secure first hand information about the chances of viability and normal development in chicks incubated under the experimental conditions described by Kuo (9, 10, 18) and Dareste (5-7).

B. Material and Methods

Three experiments were run. In the first, 90 fertile white Leghorn eggs obtained fresh from the hatchery, were set to incubate in the usual commercial type of incubator at 39.5° C. On the 3rd, 4th, 7th, 8th, and 11th days of incubation, the shell about the air space was removed from random samplings among 72 of these eggs. The number of eggs in each sampling is shown in Column 2 of Table 1. The undamaged inner-shell membranes thus exposed

TABLE 1

Results on Vaselining the Inner Shell Membrane of Incubating Chick Eggs*

Day of operation	No. of eggs used	No. dead on days: 4-7	8-12	13-17	No. hatching	Last day alive
		A. Entire membrane vaselined				
3	15	15	—	—	0	6th
4	15	14	1	—	0	8th
7	15	—	13	2	0	13th
8	21	—	15	6	0	14th
11	6	—	3	3	0	15th
		B. Membrane "spot-vaselined" over embryo				
3	30	30	—	—	0	4th
4	14	14	—	—	0	6th
Total	116	73	32	11	0	

*For results on controls, see text.

were coated with a very thin layer of melted liquid vaseline[2] kept at 39.5° C. The remaining 18 eggs served as controls. Eight unopened eggs, lying on their sides, were incubated routinely. Six eggs

[2]Tube of ordinary white vaseline, Chesebrough Mfg. Co., New York City.

had the shell removed above the air space on the 3rd day, and four others were similarly treated on the 8th day. These last 10 eggs then were placed upright on their small ends, and the exposed inner-shell membrane remained unvaselined. All the eggs were observed daily as they incubated.

In the second experiment, 44 eggs were prepared in a similar manner, but only a spot of vaseline was applied to a small localized area of the inner-shell membrane above the embryo. Thirty were done on the 3rd and 14 on the 4th day. Four other eggs incubating routinely served as controls.

In the third experiment, 37 unopened eggs were used. The entire region of the air space was painted with clear white shellac on the 1st, 3rd, 4th, 7th and 11th days. Development was watched daily by candleing. In eggs painted before the 11th day, there was a tendency for the air space to creep below the area first shellacked. In several eggs, a ring of shellac was added daily to keep up with the advancing air space. Such a procedure often led to complete covering of more than half of the egg by the 17th to 18th day. Four unshellacked eggs served as controls.

C. Results

1. *Entire Membrane Vaselined*

As will be seen in Table 1, *A*, none of the 72 chick embryos in which the entire inner-shell membrane below the air space had been vaselined survived to hatching. Of the 15 chicks prepared on the 3rd day of incubation, 10 were dead by the end of the 4th day and all were dead by the close of the 6th day of incubation. Likewise, the chicks prepared on the 4th day were dead by the end of the 8th; those vaselined on the 7th day were dead by the end of the 13th; those on the 8th, by the end of the 14th; and those on the 11th, by the 15th. Cyanosis and rapid disintegration of the vitelline vessels was the inevitable result in every case. Evaporation contributed to early death, but its effect was not as marked as in the spot-vaselined eggs. The six controls, opened on the 3rd day of incubation but not vaselined, were dead by the end of the 5th day. The four controls prepared in the same way on the 8th day were dead by end of the 11th. All eight normal, unopened controls hatched satisfactorily and developed into healthy chicks.

2. *Membrane "Spot-Vaselined" Over Embryo*

The mortality among the spot-vaselined eggs was even more rapid,

as may be seen in Table 1, *B*. In fact 10 of the 14 prepared on the 4th day were dead at the close of the 5th. Evaporation seemed to be the chief causal factor in this instance. Three of the unopened controls hatched normally, and one died of unknown cause on the 11th day.

3. *Large End Shellacked*

Tables 2 and 3 summarize the data on the 37 shellacked eggs.

TABLE 2

RESULTS ON SHELLACKING THE LARGE END OF INCUBATING CHICK EGGS*

Day shellacked	No. of eggs shellacked	No. dead on days 4-7	8-12	13-18	No. hatching (20th-21st day) Normal	Abnormal
1	2	—	1	—	—	1
3	17	7	6	3	—	1
4	4	—	—	1	1	2
7	10	—	4	4	1	1
11	4	—	—	—	2	2
Totals		37		26 dead	4	7

*For results on controls see text.

TABLE 3

ABNORMAL HATCHING RESULTING FROM SHELLACKING THE LARGE ENDS OF INCUBATING CHICK EGGS

Day shellacked	No. abnormal hatching	Anomalies	Survival after hatching
1	1	Stunted; poorly feathered; leg weakness; sits most of the time	2 days
3	1	Weak; unable to hold head erect; eyes closed	Less than 24 hours
4	2	(*a*) Abdomen enormously distended right foot clubbed; unable to stand	(*a*) 3 days
		(*b*) Distended abdomen; abnormally bent neck; unable to hold head erect; unable to stand	(*b*) Less than 24 hours
7	1	Weak and unable to hold up head	
11	2	(*a*) Weak; stunted; poorly feathered	3 days
		(*b*) Distended abdomen unable to walk	16 hours

Twenty-six of the embryos did not reach the hatching stage, the majority dying before the 15th day of incubation. Three survived until the 18th day. Of the 11 chicks which hatched, only four developed into normal healthy birds, the other seven dying shortly after hatching. These seven were abnormally developed as well; a full account of their anomalies is given in Table 3. As indicated in the section on materials and methods, the air space often tended to creep below the region of the shell first shellacked. Further shellacking led to further retreat, and sometimes more than half the egg was shellacked by the 17th day. This must have led to a severe cramping of quarters for the developing chick as well as to production of a definite asphyxia. Limbs were found pressed thin in some of the chicks dying in the late stages of development. Among the abnormal hatchings, clubbed feet, inability to carry head erect, and enormously distended abdomens were found. It is probably significant that the shellacked areas on the four eggs which gave rise to normal chicks did not extend over more than ¼ of the egg. All four of the unshellacked eggs hatched normally, and healthy chicks developed.

In the two eggs shellacked on the 1st day of incubation the allantois did not apply itself to the air space as it normally does in the course of development. Instead it attached to the side and below the region of the air space, setting up a new pocket for gaseous exchange. One chick hatched, but was deformed and lived only two days. In those eggs shellacked on the 3rd to 7th day, the allantois had already made contact with the shell membrane of the air space, but had not become applied to the rest of the membrane. As a result, the majority of the embryos died shortly after the shellac was applied; interference with gaseous exchange between the chick and the outside atmosphere is inferred. Curiously enough, three embryos of four eggs shellacked on the 4th day developed to hatching, although two were abnormal and died in the course of a few days. By the time the chicks had developed normally up to the 11th day, the vitelline vessels had, of course, applied themselves generally around the entire membrane. Therefore little or no mortality was expected when the shell about the air space was shellacked this late in development. The chances for normal, healthy hatchings after the 11th day seemed to be about 50 per cent. These findings tend to confirm those reported by Dareste in 1855.

C. Discussion

These results fail to support Kuo's implication that vaselining the inner shell membrane in no way affects the behavioral development of the chick. Every attempt was made to follow Kuo's method rigidly, especially in regard to careful technique, temperature, humidity, and other factors essential to incubation. Nevertheless, six days was the longest life of any embryo. Asphyxial conditions played a part in death in every instance.

It is apparent from the work of Romijn and Roos (19) that even in the normal incubating egg an adequate gaseous exchange between the chick and the outside atmosphere is maintained with difficulty. On the 10th day of incubation, the oxygen content within the air space is only a little less than 20 volumes per cent and the carbon dioxide is more than one volume present. Just before hatching the oxygen had been reduced almost as low as 8 volumes per cent and the carbon dioxide increased to more than 8 volumes per cent. What must conditions be between the 2nd and 4th days of incubation if the air space, with its then high oxygen content, is obliterated, and the inner-shell membrane below it coated with vaseline? This is the region of the "avian placenta," through which most of the respiratory exchange must take place.

If a state of anoxemia exists in chick embryos studied under Kuo's experimental method, an explanation for the type of behavior he has reported is forthcoming. A case in point is Kuo's observation (10) on the frequency of prehatching respiratory behavior. He observed respiratory movements before the 18th day (i.e., before the membrane was pierced preparatory to hatching) in 432 out of 620 chicks. In fact, he had found it necessary to:

> . . . Save the lives of nearly 300 chicks which would otherwise have died due to failure to break through the membranes, by simply punching with a needle, a small hole through the membrane so as to allow the chick to breathe air coming from the outside.

The findings of Romijn and Roos (19) indicate a marked physiologic anoxemia near the close of the incubation period. Hence, one might expect to find more frequent premature respiratory effort on the part of the chick embryo just prior to hatching than, perhaps, in the mammalian fetus near term. Yet it is doubtful if the phenomenon is normally so extreme as to require measures of resuscita-

tion. In spite of a normally developing state of anoxemia, the chicks under the brood hen and in the commercial incubator hatch without any outside interference.

Windle and Barcroft (21) were unable to observe initiation of respiratory behavior in the chick before the 18th day. Their observations were made on transilluminated, undisturbed eggs, as well as in eggs immediately after the inner-shell has been vaselined or removed. It is highly significant that they did find that carbon dioxide in low concentration brought about rhythmic respiratory movements even in 13-day-old chicks, and that oxygen deprivation likewise produced early respiratory efforts. Essentially the same phenomena were observed in the duck (24). In the guinea pig, as reported recently (23), and in the cat and man (20) intra-uterine respiratory movements are the exception and not the rule during normal gestation. They are occasionally seen shortly before term. However, it is quite possible to induce respiratory-like rhythms and even dyspneic gasps in cat fetuses 30 to 31 days old (term is about 67 days) by increasing the carbon dioxide or decreasing the oxygen in atmospheres breathed by the mother (22). Granting that Kuo has been more successful than others in keeping chicks alive by the vaselining technique, it would appear that he has, thereby, paved the way for the early appearance of respiratory movements of anoxial origin.

These same conditions may favor the predominance of mass behavior which has been reported for the chick (13, 16). Kuo's observations seem to lend support to that school of thought which stands for an individuation of reflex activity from an already totally integrated pattern. It has been demonstrated experimentally in the cat (22), that anoxemia will suppress precise, localized responses of the fetus in favor of more tonic, sustained mass responses. In the chick, Windle and Barcroft (21) have demonstrated that carbon dioxide and anoxemia first lead to an increase in all somatic activity, a sort of generalized response, but later there is a depression of the finer movements. Such findings run counter to Kuo's contention (16) that anoxemia increases the frequency of local reflexes. It should be recalled that by "normal" conditions Kuo means his standard method with the inner-shell membrane vaselined. These are actually not normal but asphyxial conditions. By "anoxemic" conditions he means those obtaining when ¾ of the shell and the shell membrane too were coated with vaseline. Just which part of the

shell was covered is not clearly designated. Replacing the egg-shell cap and vaselining over it would not change conditions very much. It would merely enhance an already anoxemic condition. In the present experiment the mortality was more frequent and more sudden after vaselining the membrane alone than after shellacking the large end alone. Vaselining the small end should likewise have little effect. Baudrimont and St. Ange (2) and Dareste (5) reported that varnishing the small end of the egg alone interfered with the normal development of the chick in no way, while varnishing over the air space usually resulted in asphyxial death. If the egg-shell cap were left off and the vaseline were then applied ¾ of the way down the shell, it is quite reasonable to assume that conditions would be worse than those which Kuo designated "normal." Here again it seems to be a case of going from bad to worse. And if this is true, Kuo's findings can be reconciled with the other observations cited for the cat (22) and for the chick (21).

The irritability of the fetal nervous system varies with conditions of respiratory metabolism. Irritability is heightened for a brief interval upon first increasing the carbon dioxide level or decreasing the oxygen level in the fetal blood. Under such conditions the individuality of responses is clearly defined (22). Kuo appears to have missed seeing this in his chicks for various reasons. It may be that he has not stimulated soon enough after the removal of the egg-shell cap. In addition, amniotic contractions which begin as early as the 4th day may have obscured observations on the earliest local movements even if observations were made immediately. After the 9th day when these contractions have ceased it is too late to say definitely how reflexes then present have developed. Thus an extraneous mechanical factor in the chick's environment also casts the die in favor of mass activity. But more than this, daily incubation under a vaselined membrane creates an anoxemic condition—the "normal" condition in Kuo's experiments.

In the cat, further increase in the degree of anoxemia leads to depression of irritability on the afferent side of the fetal nervous system and to a simultaneous discharge of larger and larger groups of motor neurons, either "spontaneously," through some form of direct chemical stimulation, or in response to strong afferent stimulation. Thus, one encounters mass reactions in the same specimens which showed localized responses when physiologic conditions were good (22). Here is an explanation for the predominance of mass

behavior in Kuo's chicks under his "normal" conditions. How to account for the incidence of local reflexes under Kuo's admitted "anoxemia" is another matter.

With extreme degrees of asphyxia a break down of the total response often occurs and, before complete depression of activity is invoked, the last movements to go are usually those used in forced inspiration (gasping). This may account for an apparently localized head-extension reaction.

Kuo makes much of the fact that the yolk sac and cramped quarters inhibit trunk responses, particularly during the later stages of incubation. Such conditions he claims favor local reflexes of the head and limbs. They may very well be the last agonal responses of a completely asphyxiated organism responding in toto but with certain regions of the body held in environmental restraint. Long after all neural activity has ceased in the fetus, the heart continues to beat, and skeletal muscles can be made to contract when stimulated directly by faradic shocks.

In respect to the modus operandi which determines the individuation of reflex activity from the total mass reaction, Kuo disagrees with Coghill and his associates. It is environmental interference largely, which determines the process of individuation according to Kuo (13). This principle is applicable even in the origin of the complex reflexes involved in the alternate movements of progression.

> After the 11th day, the yolk sac is over the legs of the chick, and the legs are folded on the breast. . . . Under these conditions extensor thrust of the legs is interfered with by the amnion and especially by the yolk sac which are pressing against the legs. Now in order to permit movements, one of the legs has to push up the amnion and the yolk sac, thus lessening the pressure on the other leg and allowing the latter more room for movement.

Such a statement is probably a good description of what happens, but it is hardly an adequate explanation for the genesis of a new pattern of behavior. One would like to know more about the proprioceptive pathways for the conduction of these pressure stimuli; and about the development of their necessary connections to effect a motor response. Kuo does not admit a neural interpretation, and is dissatisfied with Coghill's attempt at such an explanation for Amblystoma (4). A recent paper by Youngstrom (26) seems to

have supplied the necessary support for Coghill's thesis as far as Amblystoma is concerned. A secondary motor innervation is described. It seems to explain the process of individuation in this amphibian on an entirely neurological basis. So far, no satisfactory histological proof of a neural mechanism for a totally integrated pattern has been forthcoming in either the bird or mammal.

Kuo has long stressed the important rôle of environment in the control of behavior. The present experiment serves only to re-emphasize this importance. Experimentation usually alters the normal environmental situation in some respect. In subjecting the living organism to controlled conditions for the purpose of observation, the experimental situation easily becomes identified as the normal situation, and differences between the two are often overlooked or slighted when experimental results are evaluated. In the present case, for example, it is not normal for eggs to incubate upright on their small ends—yet the "normal" sequence of stages prerequisite to hatching (10) were based on this abnormal posture. It is no more normal for a chick to effect its respiratory exchange through a vaselined membrane than it is for a mammalian fetus to shift for itself in the water bath when its umbilical cord has been clamped. Behavior in either situation is now complicated by a further factor, anoxemia, and cannot be assumed to be of the same quality as behavior normally obtaining in ovo or utero under conditions of adequate oxygenation.

D. Summary

A critique of Kuo's position (9-18) in the present controversy over the nature of early embryonic movements has been presented. It has been suggested that interference with gaseous exchange through the inner-shell membrane in the region of the air space leads to a condition of anoxemia which must be considered in evaluating Kuo's results. Additional experimentation on vaselining this region of the inner-shell membrane, and on shellacking the shell over the air space in the incubating chick's egg have indicated the following:

1. Upon vaselining the membrane at the air space either wholly or in part, none of the embryos lived longer than six days after the operation.

2. Cyanosis and rapid disintegration of the vitelline vessels was the inevitable result in every case. Asphyxia seemed to be the chief causal factor in death, although evaporation appeared to be an im-

portant factor in the death of "spot-vaselined" eggs, and in the opened, upright, unvaselined controls.

3. Unopened controls, lying normally on their sides, hatched and developed into healthy chicks.

4. The results of shellacking the shell over the air space varied according to the time of shellacking: (*a*) Eggs shellacked on the first day of incubation before the allantois had applied itself to the inner shell membrane at the air space usually hatched because a new air pocket was set up elsewhere. Some chicks developed abnormally, however. (*b*) Eggs shellacked on the 3rd to 7th day of incubation, after the vitelline vessels had made contact with the membrane at the air space but had not become extensively attached elsewhere to the membrane, very seldom hatched. Of the few chicks that did hatch, most were malformed and died within a few days. (*c*) Eggs shellacked on the 11th day, or after the vitelline vessels had applied themselves extensively over the entire inner-shell membrane, usually hatched. But the prognosis for healthy chicks was none too good.

References

1. Angulo y González, A. W. The prenatal development of behavior in the albino rat. *J. Comp. Neur.*, 1932, **55**, 395-442.
2. Baudrimont, A., & Ange, M. St. Recherches sur les phenomenes chimiques de l'evolution embryonnaire des Oiseux et des Batrachiens. *Ann. de Chim. (et de physique)*, 1847, Series 3, **21**, 195-295.
3. Carmichael, L. Origin and prenatal growth of behavior. In *A Handbook of Child Psychology.* Worcester, Mass.: Clark Univ. Press, 1933. (Pp. 31-159.)
4. Coghill, G. E. Anatomy and the Problem of Behavior. New York: Macmillan, 1929. (Pp. 11 and 113.)
5. Dareste, M. C. Recherches sur l'influence que des enduits impermiables appliques sur la coquille de l'oeuf excercent sur le developpedu poulet. *Comp. Ren. Acad. Sci.*, 1855, **41**, 963-966.
6. ———. Memoire sur l'influence qu'exerce sur le developpement du poulet. (L'application partielle d'un vernis sur la coquille de l'oeuf). *Ann. Sci. Nat.*, 1855, *Series* (Zoologie) **4**, 119-129.
7. ———. Note sur l'influence exercee sur le developpement du poulet par l'application locale d'un vernis sur la coquille de l'oeuf. *Comp. Ren. Soc. Biol.*, 1857, **9**, 99-102.
8. Hooker, D. Early fetal activity in mammals. *Yale J. Biol. & Med.*, 1936, **8**, 579-602.
9. Kuo, Z. Y. Ontogeny of embryonic behavior in aves: I. Chronology and general nature of behavior of chick embryo. *J. Exper. Zoöl.*, 1932, **61**, 395-430.

10. ———. Ontogeny of embryonic behavior in aves: II. Mechanical factors in the various stages leading to hatching. *J. Exper. Zoöl.*, 1932, **62**, 453-483.

11. ———. Ontogeny of embryonic behavior in aves: III. Structural and environmental factors in embryonic behavior. *J. Comp. Psychol.*, 1932, **13**, 245-241.

12. ———. Ontogeny of embryonic behavior in aves: IV. Influence of embryonic movements upon behavior after hatching. *J. Comp. Psychol.*, 1932, **14**, 109-121.

13. ———. Ontogeny of embryonic behavior in aves: V. Reflex conception in light of embryonic behavior in birds. *Psychol. Rev.*, 1932, **39**, 499-515.

14. ———. Ontogeny of embryonic behavior in aves: XII. Stages in the development of physiologic activities in chick embryo (a review). *Amer. J. Psychol.*, 1933, **51**, 361-378.

15. ———. Studies in the physiology of the embryonic nervous system: I. Effect of curare on motor activity in chick embryo. *J. Exper. Zoöl.*, 1939, **82**, 371-396.

16. ———. Studies in the physiology of the embryonic nervous system: II. Experimental evidence on the controversy over the reflex theory in development. *J. Comp. Neur.*, 1939, **70**, 437-458.

17. ———. Studies in the physiology of the embryonic nervous system: III. Local reflexes or total pattern? *Psychol. Rev.*, 1939, **46**, 93-122.

18. Kuo, Z. Y., & Carmichael, L. A technique for the motion-picture recording of the development of behavior in the chick embryo. *J. of Psychol.*, 1937, **4**, 343-348.

19. Romijn, C., & Roos, J. The air space of the hen's egg and its changes during the period of incubation. *J. Physiol.*, 1938, **94**, 365-379.

20. Windle, W. F. Physiology of the Fetus. Philadelphia: Saunders, 1940.

21. Windle, W. F., & Barcroft, J. Some factors governing the initiation of respiration in the chick. *Amer. J. Physiol.*, 1938, **121**, 684-691.

22. Windle, W. F., & Becker, R. F. Relation of anoxemia to early activity in the fetal nervous system. *Arch. Neur. & Psychiat.*, 1940, **43**, 90-101.

23. Windle, W. F., Becker. R. F., Barth, E. E., & Schulz, M. D. Aspiration of amniotic fluid by the fetus: An experimental roentgenological study in the guinea pig. *Surg., Gynec. & Obst.*, 1939, **69**, 705-712.

24. Windle, W. F., & Nelson, D. Development of respiration in the duck. *Amer. J. Physiol.*, 1938, **121**, 700-707.

25. Windle, W. F., Orr, D. W., & Minear, W. L. The origin and development of reflexes in the cat during the third fetal week. *Physiol. Zoöl.*, 1934, **7**, 600-617.

26. Youngstrom, K. A. A primary and a secondary innervation in Amblystoma. *J. Comp. Neur.*, 1940, **73**, 139-151.

Department of Anatomy
Northwestern University Medical School
303 East Chicago Avenue
Chicago, Illinois

The Journal of Genetic Psychology, 1942, **60**, 167-175.

THE SOCIAL COMPETENCE OF IDENTICAL TWINS*

Department of Psychology, The University of Buffalo

EVELYN TROUP AND OLIVE P. LESTER

Psychologists, interested in the adjustment of individuals, have become increasingly aware of the significance of social development and competence. The importance of social adequacy as a criterion has long been utilized in the diagnosis of feeblemindedness, and the emphasis on this phase of behavior is now being extended to normal subjects who present problems of adjustment and to those for whom this aspect of behavior study seems desirable.

Doll (2-6) and Furfey (7-9) in particular have been interested in this problem of social adequacy and maturity and have devised scales that measure aspects of it.

The present investigation is concerned with the application of one of these scales, the Doll *Social Maturity Scale,* to identical twins reared together. The writers have accepted for purposes of this study Doll's concept of social competence, i.e., social independence and responsibility as manifested in six major categories of social behavior—self-help, locomotion, communication, occupation, self-direction, and socialization, and have proceeded to study this aspect of behavior as thus defined.

The choice of identical twins as subjects permits further investigation of similarities and differences, both qualitatively and quantitatively in these subjects as well as furnishing pertinent data on the problem of the effect of environment upon development.

A. CHOICE OF SUBJECTS

Sixteen pairs of identical twins, eight male and eight female, were rated upon the Doll *Social Maturity Scale.* All the twin pairs were either in elementary or in high school at the time of the interviews. The age range was 13 to 17 years.

The question of monozygosity of these subjects is most important.[1]

*Received in the Editorial Office on November 5, 1940.

[1]The choice of twins is based upon the general technique developed by D. Cecil Rife, Genetic studies of monozygotic twins: I. A diagnostic formula. *J. Hered.,* 1933, **24**, 333-345. This diagnostic formula consists of four quali-

It is recognized that only in the case of true identical twins is variability in the genotype avoided. Although the heredity of monovular twins is never identical, still there is more likeness than in any other pairs of organism, the degree of similarity being greater than ordinarily found among siblings. Thus, it may be assumed that hereditary factors as causes of differences are reduced to a minimum and the influence of environment may be studied more satisfactorily. The small number of subjects makes the question of monovularity of even greater importance.

B. Procedure

1. *Interviews*

Two investigators were utilized, one interviewing the mothers of each of the 16 pairs, the other the twin members of each pair. In the case of the twins themselves the interviews took place in the school, a separate room being given over for that purpose. The twins were interviewed in immediate succession leaving no opportunity for them to exchange ideas relative to the questions asked. In the interviews with the mothers, only one twin was discussed at the first interview, a second visit following after a period of a month or six weeks. This plan was carried out in order to avoid any tendency on the part of mothers to report the two members of a pair as closely alike as possible. No difficulty in gaining coöperation was experienced because of the former contacts of one of the investigators with both children and parents.

2. *Scoring the Responses*

The responses were taken verbatim and interpreted upon the bases set forth by Doll in his manual of instructions. In the case of a few doubtful replies, the experimenters agreed upon the interpretation to be placed upon them. Social quotients (*SQ*'s) were then computed by bringing scores into relation with chronological age in the fashion similar to that employed in obtaining the Intelligence Quotient (*IQ*) (4).

---

tative traits—blood groups, *M-N* agglutinogen reaction, presence or absence of hair between the first and second joints of the fingers, ability to taste phenyl-thio-carbamide and four quantitative traits—stature, iris pigmentation, intelligence quotient, and ridge count of finger patterns. This was modified where necessary to allow for limitations in this set-up. The chances that these twins are monozygotic is .999.

TABLE 1

INTELLIGENCE QUOTIENTS (*IQ*'S), SOCIAL QUOTIENTS (*SQ*'S), AND INTRA-PAIR DIFFERENCES OF SIXTEEN PAIRS OF IDENTICAL TWINS

	Boys							Girls					
Pair	*IQ*	Intra-pair *Diff.*	*SQ* Subj. Rep.	Intra-pair *Diff.*	*SQ* Moth. Rep.	Intra-pair *Diff.*	Pair	*IQ*	Intra-pair Diff.	*SQ* Subj. Rep.	Intra-pair Diff.	*SQ* Moth. Rep.	Intra-pair Diff.
I	80	0	80	2	99	0	I	98	7	109	0	108	0
	80		78		99			91		109		108	
II	99	4	100	4	100	0	II	107	6	106	0	103	0
	95		96		100			101		106		103	
III	102	0	110	0	109	0	III	99	1	113	0	106	0
	102		110		109			98		113		106	
IV	105	0	92	2	97	0	IV	88	11	74	0	96	4
	105		90		97			77		74		92	
V	100	2	107	8	105	4	V	87	9	93	0	95	2
	98		99		101			78		93		97	
VI	101	0	103	0	99	0	VI	106	6	110	2	104	0
	101		103		99			100		108		104	
VII	101	0	108	8	110	0	VII	101	0	102	0	104	0
	101		116		110			101		102		104	
VIII	101	1	107	2	101	1	VIII	104	11	92	3	94	0
	100		105		100			93		95		94	

C. Results

1. *General Findings*

Table 1 gives the complete data on the Intelligence Quotients (*IQ*'s) based upon the Stanford Revision of the Binet, Old Form, and Social Quotients (*SQ*'s) for the 16 pairs of twins, obtained from the interviews with the subjects and with the mothers.

It will be observed in Table 1 that all but three pairs of twins fall in the so-called "normal" classification according to the Stanford-Binet grouping. Also there are no twin pairs above 107 *IQ*. It might be inferred from the absence of any twins of high average or superior ability that this is not a representative group of identical twins. The difficulty of finding a fairly large group of identical twins of at least normal mentality was an interesting fact brought out in a previous study by one of the writers. The lower average mental capacity of this group as compared with other type of twins and of the unselected population is confirmed in a study by Byrns (3) in which she found the median test percentile on the Henmon Nelson test of 376 twins was more than 10 points lower than that of the general population of high school students.

It is true, however, that this group of twins represents a somewhat selected identical twin population when lower mental levels are taken in account. Here there was a definite attempt to eliminate twins below 70 *IQ*, because limited mental capacity would not permit adequate participation in a testing program. All interpretations of results are to be made in terms of this selection.

Table 2 which follows presents in somewhat greater detail the

TABLE 2

Range, Mean *IQ*'s, *SQ*'s, and Intra-Pair Differences in Sixteen Pairs of Identical Twins

	Boys	Girls	Group
Range of IQ's	80-105	77-107	77-107
Range of *SQ*'s—subject reporting	78-116	74-113	74-116
Range of *SQ*'s—mothers reporting	97-110	92-108	92-110
Mean *IQ*	98.19	95.56	
Mean *SQ*—subject reporting	100.25	99.94	
Mean *SQ*—mothers reporting	102.19	101.13	
Mean intra-pair difference in *IQ*	.875	6.375	
Mean intra-pair difference in *SQ* subject reporting	3.25	.625	
Mean intra-pair difference in *SQ* mothers reporting	.625	.750	

findings relative to IQ's, SQ's, range of scores, and intra-pair differences in the twins:

There is no significant difference in the mean IQ's of the girls and boys, though there is considerably greater variability in intelligence in the twin girl pairs than in the boys. This is revealed in a significant difference of 5.5 between the average intra-pair differences in the two groups.

The mean SQ's as determined both on the basis of own report and mothers' are within the normal range for social competence and show no significant differences between the sexes. There is a slight tendency for the SQ's with mothers reporting to be higher than similar quotients when subjects are reporting, but this difference is not significant.

Analysis of the data reveals the fact that in regard to the social quotients, boy twins, when they are reporting, show considerably greater intra-pair differences than do girls—a reversal of the data relative to IQ. However, though this sex difference is only 3.23 times the probable error of the difference, and hence does not meet the criterion for complete reliability, it nevertheless is interesting and somewhat significant to find it at all in view of the similar environments in which these subjects have grown up.

Another interesting point is the tendency which reveals itself in the average intra-pair differences and in the range of SQ's for mothers to see their twin offspring in not only a similar light, ignoring differences, but to see the intellectually inferior ones as being more socially adequate than they are—at least as measured by those subjects' own replies. See Pair I (Boys) and Pairs IV and V (Girls).

In four twin pairs (Boys II, V, VIII, Girls VI) the higher IQ twin also scores higher in SQ (own report) and this is the case in three twin pairs when mothers report (Boys V, VIII, and Girls IV). In general, however, intellectual differences do not reveal themselves in corresponding social competency scores.

One could speculate a little on this lack of variability in social maturity (SQ's) in those twin pairs who show intra-pair differences in intelligence of six or more points. Perhaps such relatively slight differences in intelligence as are found in these subjects do not reveal themselves in social adequacy differences. Possibly family and the general environmental settings enforce a uniformity on children even where differences in intelligence exist. It is also possible that the measuring instrument is not able to reveal differences which may exist between the twin pairs in social maturity.

2. *Resemblance of Identical Twins in Social Maturity as Compared with the Resemblance in Intelligence*

Table 3 shows the degree of resemblance in these identical twin pairs as revealed by correlation.[2]

TABLE 3
DEGREE OF INTRA-TWIN RESEMBLANCE IN *IQ*'S AND *SQ*'S IN SIXTEEN IDENTICAL TWIN PAIRS

	Rho	*PE*
Social Quotients, twins reporting	.94	.021
Social Quotients, mothers reporting	.98	.007
Intelligence Quotients	.77	.072

These identical twins resemble one another very markedly in social maturity, and this resemblance is greater when mothers alone report the data. Here again the tendency on the part of mothers to think and see their twin-offspring similarly seems to be apparent.

The intra-twin resemblance in intelligence is not as high as in social competence. The coefficient of correlation (+.77 *PE* ±.072) is lower than those usually reported in the literature—such *r*'s usually being in the neighborhood of +.90.[3] It is very likely that the correlation coefficients showing these resemblances have been depressed by the selective factor previously referred to. The lower levels of intellectual ability (*IQ* below 70) were purposely excluded and this resulting homogeneity in the group would have a tendency to lower the coefficients. It must be mentioned, however, that by the same token the intra-twin resemblances in social competency are also lowered.

The data present fairly clear-cut evidence that for these pairs of twins at least, environmental factors play a considerable rôle in making these twins alike in their social competence—more alike than in intelligence.

Social development, including as it does the development of social responsibility and independence (self-help, locomotion, communica-

[2]All correlation coefficients were obtained by using the method of rank differences. The number of cases, together with the fact that the range of abilities in these pairs does not conform to the normal distribution curve seemed to justify its use instead of the product moment formula.

[3]One of the present writers working with 20 pairs of identical twins of which these 16 pairs of subjects now being reported on were included, reports a coefficient of correlation for *IQ*'s of +.81±.037.

tion, self-direction and socialization), is markedly influenced by the "restraints" and "freedoms," "obligations" and "duties" set up in these similar environments of the twin pairs.

3. *Relationship of Social Competence and Intelligence*

Table 4 presents the correlations between social quotients (*SQ*'s) and Intelligence Quotients (*IQ*'s) for each member of the twin pairs.

TABLE 4

CORRELATION COEFFICIENTS BETWEEN *IQ*'S AND *SQ*'S IN SIXTEEN IDENTICAL TWIN PAIRS

	Rho	*PE*
Twins reporting	+.37	.107
Mothers reporting	+.30	.114

These data indicate no significant relationship between the social and intelligence quotients in these twin pairs. Again it may be that the elimination from the study of those twins of lower intelligence may have hidden a higher relationship existing between these two factors such as other investigators have reported, but no such relationship reveals itself in the subjects employed in this study.

In analyzing individual pairs qualitatively, interesting explanations of this low relationship between intelligence and social maturity are revealed. In one pair of boys of high average intelligence (Pair IV) personality factors (the boys are shy and withdrawn), family supervision and solicitation have functioned against the attainment of social competence as measured in this study.

One pair of girls of low average intelligence (Pair III) developed maximum independence and self reliance due to the father's death which resulted in the mother's employment outside the home which threw household responsibility very much upon them.

Pair V (Girls) reveals a somewhat similar picture in that insistence on self-help on the part of the mother has given these twins a social quotient somewhat above the intelligence level they possess.

Pair VII (Boys) shows average ability intellectually, but growing up in a stimulating environment which encourages considerable freedom of action, has resulted in a higher level of social competence than would be expected.

4. *Data on Reliability of Rater's Reports*

In evaluating the usefulness of the scale one important question

arises as to the consistency of scores made by two different raters. In this case there were not only two different raters, but the information was obtained from two different individuals, i.e., the subject himself and mother. The correlation coefficient between the social quotients thus obtained was $+.84\pm.037$. This high degree of relationship suggests that the results of an interview with the subjects themselves reporting can be considered a reliable estimate of social competence as measured by the Doll *Scale* and that furthermore two different experienced raters can obtain independently of one another consistent ratings from subjects and mothers.

D. Summary

In this investigation of social competence employing the Doll *Social Maturity Scale* with 16 pairs of identical twins, eight girl and eight boy pairs, reared together, the following results seem worthy of emphasis.

1. Identical twins reared together were found to resemble each other to a greater degree in social competence than in intelligenc. The resemblance in *SQ*'s expressed in terms of R is $.94\pm.021$ with subjects reporting and $.98\pm.007$ with mothers reporting. The intra-twin correlation for intelligence is $.77\pm.072$.

2. There appears to be a slight tendency for mothers to report not only a greater uniformity of behavior in the twin pairs, but also slightly higher average *SQ*'s result from their reports than from subjects' own reports.

3. Boy twins, when reporting, show greater variability in *SQ*'s (intra-pair differences) than do girl pairs, and this in the face of a significant mean intra-pair difference in intelligence in the girls which might ordinarily be presumed to make for greater differences in social competence in twin pairs.

4. In this group of identical twins there is a low positive correlation between intelligence as measured by the Stanford Revision of the Binet Test, Old Form, and the *SQ*'s derived from the Doll Scale ($R = .37\pm.107$ subjects reporting, and $R = .30\pm.114$, mothers reporting). There is the possibility that certain selective factors may account for this low relationship.

5. There is a high relationship between *SQ*'s derived from twins' own reports and mothers' reports even when separate investigators contracted these two groups ($R = +.84\pm.037$).

6. There is no evidence of sex differences in average social competency (*SQ*'s) in these pairs of twins.

References

1. Byrns, R. Mental ability of twins. *Sch. & Soc.*, 1934, **40**, 671-672.
2. Doll, E. A. Vineland social maturity scale. *Train Sch. Bull.*, 1935, **32**, 1-7.
3. ———. The measurement of social competence. *Proc. Amer. Asso. Ment. Def.*, 1935, **40**, 103-123.
4. ———. Vineland Social Maturity Scale: Revised Condensed Manual of Directions. Vineland, N. J.: Training School, 1936.
5. ———. A practical method for the measurement of social competence. *Eug. Rev.*, 1937, **29**, 197-200.
6. ———. The inheritance of social competence. *J. Hered.*, 1937, **28**, 152-165.
7. Furfey, P. H. The measurement of developmental age. *Cath. Univ. Amer., Educ. Res. Bull.*, 1928, **2**, No. 10.
8. ———. A revised scale for measuring developmental age in boys. *Child. Devel.*, 1931, **2**, 102-114.
9. ———. Developmental age. In *Toward Understanding Children. Univ. Iowa Exten. Bull.*, 1932, No. 283, 9-21.

Department of Psychology
The University of Buffalo
Buffalo, New York

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1942, 60, 177-182.

A NOTE ON THE RÔLE OF VISION IN THE FEEDING BEHAVIOR OF WHITE RATS*

Department of Psychology, University of Arkansas

STANFORD C. ERICKSEN

A. STATEMENT OF PROBLEM

In analyzing the hunger drive chief interest has been focused on its initiating and accompanying psycho-physiological conditions. Bash (1) has emphasized the integrative nature of the hunger drive. His results show that we cannot simplify the initiating factors in feeding behavior by saying that it is the result of stomach contractions alone or the like. Harlow (3) also suggested the possible influence of other than interoceptive stimulation in directing the feeding activities of rats. It would thus seem that a complete understanding of the hunger drive would require an analysis of the: (*a*) initiating conditions; (*b*) the mechanisms that continue the feeding; and, (*c*) the factors which cause the organism to stop eating. The study of the influence of some factor common to these three phases of feeding behavior may be a worthwhile approach to the understanding of the total and integrated act of feeding. Vision is one such possible factor and serves as the problem of the present experiment.[1] We are simply interested in comparing normal with blind rats in the amount and manner of eating.

B. PROCEDURE

A group of 20 male white rats about four months of age were fed individually once a day for three days. They were then matched

*Received in the Editorial Office on October 9, 1940.

[1]This is a preliminary and exploratory experiment. It is our plan to make further studies on the rôle of cutaneous sensitivity, audition, olfaction, etc., and the influence of cerebral lesions on simple feeding behavior.

Research Paper No. 706, Journal Series, University.

TABLE 1
SEQUENCE OF CONDITIONS AND RESULTS

Sequence	Conditions of feeding	Time of feeding	Number of days	Size of food can	Average daily gain in weight: Blind	Normal
A	alone	15 min.	19	small	23.5 gr.	24.7 gr.
B	two together	15 min.	9	small	25.2	26.4
C	together	20 min.	7	small	26.2	27.4
D	alone	20 min.	7	Rats Nos. 1- 5 large	24.6	25.8
				6-10 small	24.7	25.1
E	alone	20 min.	7	Rats Nos. 1- 5 small	25.0	26.2
				6-10 large	26.4	25.9
F	alone	30 min.	7	Rats Nos. 1- 5 large	28.2	29.3
				6-10 small	29.7	28.8
G	alone	30 min.	7	Rats Nos. 1- 5 small	30.4	29.8
				6-10 large	31.3	31.4
H	one blind one normal	20 min.	7	small	28.6	27.5

for weight and divided into two groups of 10. The rats in Group II were blinded by enucleation after feeding on the third day. Throughout the experiment a record was kept of the weight of each animal before and after feeding. Presumably the increase in weight after feeding represents the amount of food eaten. The animals were fed about the same time each day on a diet of dry food (Ruch, 4) mixed with a constant proportion of whole milk to give a moist paste. The rats were fed in clean metal cages, 12 x 18 x 14 inches. The amount of food in the food cups was always more than enough for each animal. The small food cups were one inch deep and two inches in diameter. In some cases larger food cups were used, two inches deep and four inches in diameter. Every attempt was made to handle both groups in the same way and each followed the same routine, the experimental variable being simply that one group was blind, the other normal. At the beginning of the experiment the average weight for the animals in Group I was 226 grams; for Group II, blind, 223 grams. Table 1 gives the sequence of the various conditions of the experiment and the results.

C. Results

These results are presented in graphic form in Figure 1. The

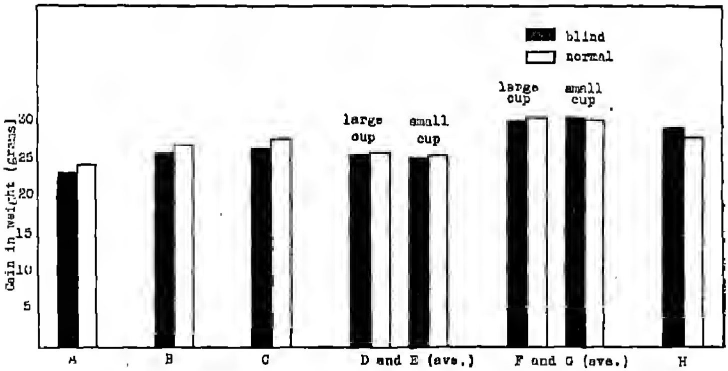

FIGURE 1

Average Gain in Weight for the Various Experimental Conditions

data in Table 1 show no statistically significant difference between the results of the blind and the normal rats. Nor does there appear to be any marked difference in the graphic record shown in Figure 1.

Conclusions to be drawn from the successive stages of the experiment are as follows:

1. *Condition A*

After 19 days of comparable feeding conditions the blind animals eat just as much and make just as large gain in weight as the normal. Within the conditions of our experiment, vision is either unimportant or is masked by other factors. In this latter connection the time limit of feeding is important. If it is too long even handicapped blind animals will have had time to eat sufficient to satisfy their present organic condition. In such a case the amount eaten is a function of such factors as size, rate of growth, health, adjustment to the conditions of feeding, etc. If the time limit is too short the lack of vision becomes largely a mechanical handicap in finding the food cup, etc.

2. *Condition B*

Our results do not show as clearly as those of Harlow (3) the presence of the "social facilitating" factor. The animals do eat a little more when feeding together in the same cage and from the same food cup than when eating alone. But there is no apparent difference between the normal and the blind animals in this respect.

3. *Condition C*

Increasing the length of the feeding period by five minutes increases the absolute amount of food eaten but shows no differentiation between the normal and blind animals. After the first 10 minutes the rate of eating falls off. It was thought that by extending the feeding time the "social facilitating" factor might be more clearly differentiated if it were present.

4. *Conditions D, E, F, and G*

There was no consistent difference in the amount of food eaten from the large or the small food cups. This result is the same for both the normal and the blind animals. Since the amount of food in either of the cups was entirely adequate for the needs of the animals there would be no biological purpose in responding differentially to them even though the rats sensed the difference in the size of the food object. Yoshioka (5) has shown the importance of this interpretation.

5. *Condition H*

Animals of approximately the same size from each of the two groups were paired together in the feeding cages. The amount of food eaten was about the same for both groups; there did not appear to be any dominance of the normal rats over the blind or the reverse. For the first two or three days there was some fighting and "incompatibility" in the feeding cages. Since the time of feeding was reduced to 20 minutes the gain in weight also dropped for both groups.

It may be noted from Figure 1 that the normals consistently gain a little more than the blind rats. This is not, however, a significant difference.

D. Discussion

In this experiment blind animals were compared to normals in eight different feeding conditions but in none of these is there any statistically significant or subjectively observed difference between the two groups. Evidently within the conditions of this experiment vision is not an important factor in determining the nature and extent of feeding.

It is the opinion of the writer that if this specific question of the sensory control of feeding behavior were investigated as extensively as the problem of the sensory control of maze learning, we would find very comparable results. Some of the specific senses are more important than others in controlling the behavior in question but the relative importance of any given sense quality is a function of the particular test conditions. In general, a normal and intact animal is better able to meet the exigencies of the experimental conditions than those animals with some one or more sense modality lacking. Organic sensitivity plays a particularly important and unique rôle in feeding behavior. But it is not the *sine qua non*. As investigations proceed it will be found that certain other controlling centers, cortical, subcortical, glandular, sensory, etc., are likewise important and effective in feeding behavior just as they are in maze learning, sex behavior (Beech, 2), and other particular types of activities.

E. Summary

Ten normal and 10 blind rats were compared as to the amount of food eaten during a given period each day. Eight different conditions were included in the experiment (feeding alone; two at a

time; 15, 20, and 30 minute periods; large food amounts vs. small food amounts; and blind and normal rats together). In no case was there a consistent or marked difference in the amount of weight increase after eating between the normal and blind animals.

The opinion is offered that the sensory control of feeding behavior is psychologically not very different from the results found in the problem of the sensory control of maze learning. At least there is probably no single sense that controls feeding (as we once thought kinesthesis did maze learning) but each sense is functional and operates to a greater or less extent depending on the particular conditions in which feeding occurs.

References

1. Bash, K. W. An investigation into a possible organic basis for the hunger drive. *J. Comp. Psychol.*, 1939, **28**, 109-135.

   ———. Contribution to a theory of the hunger drive. *J. Comp. Psychol.*, 1939, **28**, 137-160.

2. Beach, F. A., Jr. The neural basis of innate behavior: I. Effects of cortical lesions upon the maternal behavior pattern in the rat. *J. Comp. Psychol.*, 1937, **24**, 393-440.

3. Harlow, H. F. Social facilitation of feeding in the albino rat. *J. Genet. Psychol.*, 1932, **41**, 211-221.

4. Ruch, F. L. Experimental studies of the factors influencing the difficulty of blind alleys in linear mazes: I. Experiments with the maze patterns: *RLRLLRLRRLRL* and *LRLRRLRLLRLR*. *J. Comp. Psychol.*, 1935, **20**, 21-34.

5. Yoshioka, J. G. Size preference of wild rats. *J. Genet. Psychol.*, 1930, **37**, 159-162.

   ———. Size preference of albino rats. *J. Genet. Psychol.*, 1930, **37**, 427-430.

Department of Psychology
University of Arkansas
Fayetteville, Arkansas

The Journal of Genetic Psychology, 1942, **60**, 183-188.

A COMPARISON OF TEACHERS' DIAGNOSES OF MALADJUSTED CHILDREN WITH CLINICAL FINDINGS*

Department of Psychology, Northwestern University

George D. Lovell and Helen D. Sargent[1]

There has been much said and much written of late about "the whole child," about treating the school boy and girl each as a personality, not as merely another pupil to whom facts and techniques are handed, in the hope that they will somehow be assimilated. This emphasis on the unity of the child's experiences and upon the necessity for providing a complete environmental background appropriate to the learning process has been accepted by educators and psychologists alike. Attempts are continuously being made to find the best schoolroom atmosphere and the most favorable teaching techniques to accomplish this expansive purpose. Great emphasis is placed on individualizing the curriculum to meet the needs of the child.

An integrated personality depends on the child's learning habit patterns that are both adjustive to him and which meet with the approval of those with whom he comes in contact in any real way, which thus comply with the social norms. Gaining these habits by a process of random trial and error would be a very expensive and not too successful one. There must be a leader, a guide. One of these directors is the school room teacher. Upon her shoulders falls the responsibility first of determining what sort of treatment, motivation, etc., each child needs, and second of providing opportunity for this treatment, motivation, etc.

The first responsibility is of great importance. The child comes to school with certain habit patterns already established, with fixed, but modifiable ways of reacting to certain classes of situations. In the schoolroom he is placed with other children who have different modes of reacting and with a teacher who must be somewhat of a new authority to him. He must either use the existing response mechanisms or some that are compatible with them. Gradually new

*Received in the Editorial Office on October 26, 1940.

[1]Appreciation is expressed to Dr. John J. B. Morgan, director of the Northwestern Psychological Clinic for the use of the data for this study.

ones are built, some acceptable and some not so acceptable. The teacher must be able to find those which are not acceptable, ferret out their causes, and initiate a program of reeducation to overcome them.

How well is the teacher performing this function? A limited answer to this question has been found in a survey of cases referred to the Northwestern University Psychological Clinic from schools of Evanston, Illinois, and surrounding communities. When a teacher refers a child to the clinic, she fills out a referral blank which suggests the reasons for reference. From these blanks can be determined the chief complaint which can be compared with the findings both of tests and of the examiner's interviews with the child. Such a study was made on a group of 370 male cases referred during the past four years. Most of these cases were below the eighth grade at the time of referral, but a few high school pupils were included.

The data were treated in a purely statistical manner. No attempt was made to follow any one case as to reason for reference and diagnosis, but the number of cases in each category was computed in relation to the total group of cases. While there was only one major reason for reference[2] for each child, there may have been several diagnoses; i.e., difficulty may have been found in several areas. There was always a Stanford-Binet *IQ* along with diagnoses as to behavior, attitudes, and other causal agents.

Not only are reasons for reference and diagnoses by examiners given in the figures below, but where applicable and on those cases where scores are available, the number of children scoring high on the different sections of the Rogers *Personality Adjustment Scale for Children* are quoted. The number of children scoring high on the Rogers and the number of children so diagnosed will not necessarily agree. There may be two explanations for this fact. First, in a small number of the cases, other factors may outweigh the test results in the mind of the examiner. In many more cases, however, the high score on the Rogers is merely a contributing factor to a more pertinent problem, which problem is the one that receives the chief diagnosis of the examiner.

The Rogers scale is divided into four scoring sections, personal inferiority, social maladjustment, family maladjustment, and day-dreaming, with a total score for personality maladjustment in general.

[2]Actually the teacher may have checked more than one difficulty, but she placed most emphasis on the one basic reason.

In the writers' opinion, this is one of the better personality tests for children because of its interest to the child and because of its construction which permits getting important information that the child could not give directly. While the scores are not absolutely accurate, as neither are the scores of any such scale, they are definite indicators of areas which need careful attention, if the child is to be helped in his difficulty.

The available data, punched on Hollerith cards, made it possible to check each reason for reference against the actual diagnoses made for the cases having such reason for reference. Out of 370 male cases, 256 of them were represented in the four reasons for reference: mental rating, placement, poor school work, and social maladjustment, even though some 28 other reasons for reference were available. Since personality measures are not always taken when the request is purely for placement or mental rating, these two reasons for reference were not broken down as to diagnosis. Many of these diagnoses were mainly in terms of *IQ* and achievement test scores. Referrals for poor school work were, however, compared with diagnoses. This comparison appears in Table 1.

TABLE 1

EXAMINERS' DIAGNOSES OF 129 CASES REFERRED FOR POOR SCHOOL WORK

(Each case may be represented by more than one diagnosis, and any diagnosis with less than 10 cases was disregarded.)

	No. cases
Family maladjustment	63
Below *IQ* of 90	38
Feelings of inferiority	33
Social maladjustment	21
Physical disability	21
Daydreaming	12

It is rather clear that since only 38 of the 129 cases in this category had *IQ*'s below 90, and that many other factors were pertinent, the teacher was wise in sending the child to the clinic for observation. Her chief diagnosis of "poor school work" seemed to be due not to a lack of ability but to other causal agents.

As Table 1 indicates, these causal agents had to do largely with the personality adjustment of the child. Ranking high among these was family maladjustment. This study adds further evidence to the many investigations which have shown conditions in the home to have

a direct bearing on the child's behavior and work in the schoolroom. A careful teacher can often recognize this source of difficulty and take steps to minimize it or to lessen its effect. More insidious and hard to ferret out is feeling of inferiority which is the second largest contributing factor to poor school work (after low *IQ* is excluded). Social maladjustment is more easily recognized as should be physical disabilities, but the last large cause, daydreaming, requires careful attention to recognize and discriminative judgment to find the cause for it. It is not implied that these maladjustments exist in isolated form; they can and do often occur together, and such is recognized in Table 1 which includes more than one diagnosis for several of the 129 cases referred for poor school work. There were also a few scattered diagnoses other than the ones shown in Table 1 which were not included in this study because of their small number.

An observation of the diagnosis side of the picture made clear that, although there were many possible diagnoses, certain ones were predominant, both in the case of referrals for poor school work and for other reasons as well. The chief ones of these were feelings of inferiority, social maladjustment (fighting, bullying, showing off, dishonesty); physical disabilities (visual, hearing, glandular, general physical weakness, crippled, etc.); family maladjustment (fixations, family discord, low economic status, overdiscipline, inconsistency, over-stimulation, neglect); and daydreaming (excessive imagination, phantasy lying, imaginary comparisons). Table 2 indicates the number of times each of these diagnoses was made by an examiner, with the number of cases referred by the teacher for that reason shown beside it. Rogers Personality indications of maladjustment in

TABLE 2

RELATION BETWEEN CLINICAL FINDINGS (EXAMINERS' DIAGNOSES) AND THE REASONS FOR REFERENCE OF TEACHERS

	No. referred by teacher	No. scored high by Rogers	No. diagnosed by examiner
Feelings of inferiority	3	79	86
Family maladjustment	6	86*	157**
Social maladjustment	34	131	65
Physical disability	2	—	51
Daydreaming	3	84	28

*Does not include as many factors as the examiners' diagnosis of family maladjustment.

**Includes fixation, overindulgence, neglect, overstimulation, family discord, inconsistent treatment.

these areas are also shown, except for physical disability. That the Rogers and the diagnoses do not agree absolutely has been referred to above with reasons suggested.

It can readily be seen that those areas which are the most important in the eyes of the examiners of these children have not been so regarded by the teacher who referred the child. This does not say that she has not observed these tendencies in the child, but that she has considered some other thing as more important and more pertinent. It is interesting to note that teachers are more aware of social maladjustment than of the other difficulties, probably because it is more overt and readily recognizable. Daydreaming, feelings of inferiority, and difficulty caused by family maladjustment may be more easily mistaken for other things. The quiet child, and even the compensating child are less easily detected as having feelings of inferiority, or of being neglected at home. The reasons for his poor school work, reading difficulties, boasting, etc., are not so easily ferreted out. They require careful checking and knowledge of what areas might be concerned. That physical disabilities are overlooked as major causes for difficulty so easily is not understandable in many cases. It seems that with the emphasis on sound physical bodies, this area would be suspected by teachers before any others.

These findings suggest definite implications for teaching. Whenever the services of a clinic are available, the teacher shows foresight who sends the child there for observation, if she has any doubt as to her own "diagnosis." The psychological clinic can play a big part in supplementing the teacher's treatment of her problems.

Where there is no psychological clinic, with its trained personnel, the teacher can do well to note the major areas of maladjustment in children as brought out in this study and to investigate them to the best of her ability when a child seems to need some sort of help. In fact, all teachers would do well to keep open minds on deciding a child's difficulties by checking these possibilities in each case, and thus by a process of elimination, if by no other, come to a more nearly correct solution than she would otherwise. "Individualizing the curriculum" will then take on more meaning, and the child will be understood as a personality and so treated and motivated.

Suggestions for the treatment of the above main maladjustments in children are not included in this paper because of the many existing articles and books on child psychology which deal with these very problems.

SUMMARY

From a study of 370 male cases (largely below the eighth grade) referred to the Northwestern University Psychological Clinic, it was found that most of them were referred for mental rating, placement, poor school work, and social maladjustment. Only 38 out of 129 referred for poor school work were found deficient in *IQ*. The most important of the other diagnoses were physical disabilities, feelings of inferiority, social maladjustment, family maladjustment, and daydreaming. These were also the most frequent diagnoses for other reasons for reference, even though they were rarely listed, themselves, as reasons for reference. It was suggested that a psychological clinic plays an important part in supplementing the teacher's finding areas of difficulty in her pupils. But where psychological help is not possible, this study indicates the areas which are most usually annoying, and which the teacher may check in each pupil as possible sources of maladjustment which she can try to correct, thus "individualizing the curriculum" in a more meaningful manner and promoting an integrated personality.

Department of Psychology
Northwestern University
Evanston, Illinois

The Journal of Genetic Psychology, 1942, 60, 189-194.

BOOKS

The *Journal of Genetic Psychology*, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than $2 per printed page and not more tran $3 per printed page.

Conditions. Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (*a*) Makes an important theoretical contribution; (*b*) consists largely of original experimental research; (*c*) has a creative or revolutionary influence in some special field or the entire field of psychology; (*d*) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. Just send in your review. It does not matter if the book in question has been reviewed before.

(1939)

ROSETT, J. The Mechanism of Thought, Imagery, and Hallucination. New York: Columbia Univ. Press, 1939. Pp. 289.

PRATT, C. C. The Logic of Modern Psychology. New York: Macmillan, 1939. Pp. 185.

DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., SEARS, R. R., *et al.* Frustration and Aggression. New Haven: Yale Univ. Press, 1939. Pp. 209.

GESELL, A., CASTNER, B. M., THOMPSON, H., & AMATRUDA, C. S. Biographies of Child Development. New York: Harper, 1939. Pp. 328.

GOLDSTEIN, K. The Organism. New York: World Book, 1939. Pp. 533.

BARTLETT, F. C., GINSBERG, M., LINDGREN, E. J., & THOULESS, R. H. The Study of Society. New York: Macmillan, 1939. Pp. 498.

MAIER, N. R. F. Studies of Abnormal Behavior in the Rat. New York: Harper, 1939. Pp. 81.

WECHSLER, D. The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, 1939. Pp. 229.

WOODWORTH, R. S. Psychological Issues. New York: Columbia Univ. Press, 1939. Pp. 421.

KARDINER, A. The Individual and His Society. New York: Columbia Univ. Press, 1939. Pp. 503.

STEVENS, B. The Psychology of Physics. Manchester, England: Sherratt & Hughes, 1939. Pp. 282.

ELLIS, H. My Life: Autobiography. New York: Houghton-Mifflin, 1939.

(1940)

CANTRIL, H. The Invasion from Mars: A Study in the Psychology of Panic. Princeton: Princeton Univ. Press, 1940. Pp. 228.

RASHEVSKY, N. Advances and Applications of Mathematical Biology. Chicago: Univ. Chicago Press, 1940. Pp. 214.

BURLINGAME, L. L. Heredity and Social Problems. New York: McGraw-Hill, 1940. Pp. 369.

DENNIS, W. The Hopi Child. New York: D. Appleton-Century, 1940. Pp. 204.

HILGARD, E. R., & MARQUIS, D. G. Conditioning and Learning. New York: Appleton-Century, 1940. Pp. 429.

HARTMAN, G. W., NEWCOMB, T., *Editors*. Industrial Conflict: A Psychological Interpretation. New York: Cordon, 1940. Pp. 583.

DOOB, L. W. The Plans of Men. New Haven: Yale Univ. Press, 1940. Pp. 411.

HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. B. Mathematico-Deductive Theory of Rôte Learning. New Haven: Yale Univ. Press, 1940. Pp. 329.

LANDIS, C., *et al.* Sex in Development. New York: Hoeber, 1940. Pp. 329.

THORNDIKE, E. L. Human Nature and the Social Order. New York: Macmillan, 1940. Pp. 1019.

GESELL, A., *et al.* The First Five Years of Life. New York: Harper, 1940. Pp. 393.

YOUNG, K. Personality and Problems of Adjustment. New York: Crofts. Pp. 868.

BROWN, J. F. The Psychodynamics of Abnormal Behavior. New York: McGraw-Hill, 1940. Pp. 484.

GOLDSTEIN, K. Human Nature. Cambridge: Harv. Univ. Press, 1940. Pp. 258.

KATONA, G. Organizing and Memorizing. New York: Columbia Univ. Press, 1940. Pp. 318.

BAKKE, E. W. The Unemployed Worker. New Haven: Yale Univ. Press, 1940. Pp. 465.

(1941)

ZIPF, G. K. National Unity and Disunity. Bloomington, Ind.: Principia Press, 1941. Pp. 408.

REIK, T. Masochism in Modern Man. New York: Farrar & Rinehart, 1941. Pp. 439.

BRITT, S. H. Social Psychology of Modern Life. New York: Farrar & Rinehart, 1941. Pp. 562.

BARTLEY, S. H. Vision: A Study of Its Basis. New York: Van Nostrand, 1941. Pp. 350.

CANTRIL, H. The Psychology of Social Movements. New York: Wiley, 1941. Pp. 274.

MILLER, N. E., & DOLLARD, J. Social Learning and Imitation. New Haven: Yale Univ. Press, 1941. Pp. 341.

DEARBORN, W. F., & ROTHNEY, J. W. M. Predicting the Child's Development. Cambridge: Sci-Art, 1941. Pp. 360.

FARAGO, L. (*Ed.*) German Psychological Warfare. New York: Committee for National Morale, 1941. Pp. 133.

CRITICAL REVIEWS OF RECENT BOOKS

(Bradley, Charles, M.D. Schizophrenia in Childhood. New York: Macmillan, 1941. Pp. 152.)

Reviewed by Livingston Welch

Though most psychiatry textbooks have dealt with Schizophrenia in childhood, the world has been waiting a long time for an extensive survey of this topic so that the many fragmentary reports and theories that have accumulated, may be presented in some meaningful whole. The organization of Dr. Bradley's book is admirable and the objectivity of his treatment leaves nothing further to be desired.

He begins wisely by defining dementia praecox, schizophrenia, and paregasia, thus reclarifying the different connotations of the terms. He also discusses the meaning of "dementia praecocissima," and "dementia infantilis." The former he regards as identical with childhood schizophrenia, while the latter he says "appears to be a symptom which may signalize the onset of acute schizophrenia at an early age level." He presents the following as the major features of the diagnostic criteria for dementia infantilis: (*a*) Normal development until the third or fourth year. (*b*) At this age an acute behavior change takes place characterized mainly by motor restlessness, speech disturbance, anxiety, and failure of educational progress. (*c*) Retention of an intelligent facial expression. (*d*) No signs of organic brain damage and (*e*) a very rapid mental deterioration leading to dementia.

Dr. Bradley shows the many respects in which the symptoms of children with schizophrenia differ from those of adults. In general they are fewer and simpler in children than in adults. An analysis of many reports seems to point to several major types in childhood. (*a*) Diminished interest in the environment. This appears to be most outstanding of all. (*b*) Emotional disturbances which include marked degrees of anxiety, negativism, and irritability. (*c*) Symptoms of regression or a return to a more immature and simpler levels of interest. (*d*) Alterations of motor behavior, including sudden displays of peculiar conduct, sluggish awkward gait and catatonic manifestations. (*e*) Speech disturbances of thinking manifested by a bizarre thinking process. (*f*) Hallucinations and delusions.

For a positive diagnosis the author suggests that the following items be considered: (*a*) The child must be psychotic. (*b*) His mental disorder must have appeared without known or obvious cause after a period in earlier life, when he was comparatively free from mental disorder. (*c*) He must give positive evidence of severely disturbed social contact with and interest in his surroundings. (*d*) He may show a variety of often very dramatic symptoms which may be considered as methods by which he expresses his disturbed contact with his surroundings. (*e*) He must show some evidence of regression or deterioration in his behavior. (*f*) An hereditary taint of schizophrenia favors a similar diagnosis in the child. (*g*) An entire absence of physical or neurological signs is compatible with a diagnosis of childhood schizophrenia, but this does not preclude the disorder. (*h*) A diagnosis should be made only on a basis of history, development of symptoms, and never exclusively on the basis of the resemblance of the behavior to that which is seen in adult schizophrenia.

The classical four types of schizophrenia are hardly ever distinguishable in childhood. The only justifiable divisions, according to the author are the "acute" and "chronic" forms. "More descriptive terms," he adds, "would be a periodically agitated type for those patients having an acute course and an insidious type for those whose development is chronic in nature."

So far, no conclusive studies of shock therapy in connection with childhood schizophrenia have been reported. Dr. Bradley refers to several cases in which insulin and metrazol shock treatment were used without any success. He also describes the use of sodium amytal, caffeine citrate, ephedrine sulfate, and amphetamine sulfate. The last was by far the most successful, but the author explains that "superficial symptoms rather than fundamental mechanisms are altered." As far as prognosis is concerned in childhood schizophrenia, he says, "all reports are uniformly gloomy regardless of treatment," in fact decidedly worse than in adult patients.

Institute for Research in Child Psychology
Hunter College
New York City

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i.e., the most recently received books will be found at the end of the list.)

TORREY, T. Wisdom for Widows. New York: Dutton, 1941. Pp. 252.

GOLDSTEIN, H. Reading and Listening Comprehension. New York: Bur. Publ., Teach. Coll., Columbia Univ., 1940. Pp. 69.

ZIPF, G. K. National Unity and Disunity. Bloomington, Ind.: Principia Press, 1941. Pp. 408.

MASLOW, A. H., & MITTELMANN, B. Principles of Abnormal Psychology New York: Harper, 1941. Pp. 638.

KANDEL, I. L. Professional Aptitude Tests in Medicine, Law, and Engineering. New York: Bur. Publ., Teach. Coll., Columbia Univ.; 1940. Pp. 78.

WOOD, L. The Analysis of Knowledge. Princeton: Princeton Univ. Press, 1941. Pp. 263.

COLSON, E. M. An Analysis of the Specific References to Negroes in Selected Curricula for the Education of Teachers. New York: Bur. Publ., Teach. Coll., Columbia Univ., 1940. Pp. 178.

ALBERT, A. A. Introduction to Algebraic Theories. Chicago: Univ. Chicago Press, 1941. Pp. 137.

KRAINES, S. H. The Therapy of the Neuroses and Psychoses. Philadelphia: Lea & Febiger, 1941. Pp. 512.

SKINNER, C. E., & HARRIMAN, P. L., *Editors.* Child Psychology. New York: Macmillan, 1941. Pp. 522.

CATTEL, P. The Measurement of Intelligence of Infants and Young Children. New York: Psychological Corporation, 1940. Pp. 274.

ALEXANDER, F. MATTHIAS. Man's Supreme Inheritance. New York: Dutton, 1941. Pp. 354.

SCHORER, MARK. The Hermit Place. New York: Random House, 1941. Pp. 313.

MORT, P. R., & CORNELL, F. G. American Schools in Transition. New York: Columbia University, 1941. Pp. 546.

WOOD, L. The Analysis of Knowledge. Princeton: Princeton Univ. Press, 1941. Pp. 263.

BABCOCK, H. Time and the Mind. Cambridge: Sci-Art, 1941. Pp. 304.

WOODCOCK, L. P. Life and Ways of the Two-Year-Old. New York: Dutton, 1941. Pp. 267.

CAMERON, D. E. Objective and Experimental Psychiatry. New York: Macmillan, 1941. Pp. 390.

BARTLEY, S. H. Vision: A Study of Its Basis. New York: Van Nostrand, 1941. Pp. 350.

PORTER, J., *et al.* Studies in Arithmetic. London: Univ. London Press, 1941. Pp. 218.

MAYER, J. Social Science Principles in the Light of Scientific Method. Durham, N. C.: Duke Univ. Press, 1941. Pp. 573.

OLDFIELD, R. C. The Psychology of the Interview. London: Methuen, 1941. Pp. 144.

SHERMAN, M. Basic Problems of Behavior. New York: Longmans, 1941. Pp. 440.

ANDERSON, S. E. Where God and Science Meet. Boston: Meador, 1941. Pp. 235.

BRITT, S. H. Social Psychology of Modern Life. New York: Farrar & Rinehart, 1941. Pp. 562.

VALENTINE, W. L. Experimental Foundations of General Psychology. New York: Farrar & Rinehart, 1941. Pp. 432.

MILLER, N. E., & DOLLARD, J. Social Learning and Imitation. New Haven: Yale Univ. Press, 1941. Pp. 341.

LLOYD-JONES, E., & FEDDER, R. Coming of Age. New York: Whittlesey House, 1941. Pp. 280.

GUENTHER, L. A German Ace Tells Why. Boston: *Sci-Art*, 1941. Pp. 78.

PROCEEDINGS OF AMERICAN PHILOSOPHICAL SOCIETY. Symposium on Recent Advances in Psychology. Philadelphia: American Philosophical Society, 1941.

McKINNEY, F. Psychology of Personal Adjustment. New York: Wiley, 1941. Pp. 636.

GREENE, E. B. Measurement of Human Behavior. New York: Odyssey, 1941. Pp. 777.

CASE, V. Your Personality: Introvert or Extravert. New York: Macmillan, 1941. Pp. 277.

CANTRIL, H. The Psychology of Social Movements. New York: Wiley, 1941. Pp. 274.

ROETHLISBERGER, F. J. Management and Morale. Cambridge: Harvard Univ. Press, 1941. Pp. 194.

OPLER, M. E. An Apache Life-Way. Chicago: Univ. Chicago Press, 1941. Pp. 500.

ALEXANDER, F. M. The Universal Constant in Livnig. New York: Dutton, 1941. Pp. 270.

HART, H. New Gateways to Creative Living. New York: Abingdon, 1941. Pp. 197.

ROBINSON, F. P. Diagnostic and Remedial Techniques for Effective Study. New York: Harper, 1941. Pp. 318.

WANG, C. K. A. An Annotated Bibliography of Mental Tests and Scales. Peiping: Catholic Univ. Press, 1940. Pp. 698.

MORGAN, J. J. B. Workbook in Psychology. New York: Farrar & Rinehart, 1941. Pp. 122.

FALK, H. A. Corporal Punishment. New York: Teach. Coll., Columbia Univ., 1941. Pp. 162.

TERMAN, L. M., & HOWARD, L. S. The Road Ahead. New York: World Book, 1941. Pp. 402.

HAYES, S. P. Contributions to a Psychology of Blindness. New York: Amer. Found. for the Blind, 1941. Pp. 296.

LENDE, H. Books About the Blind. New York: Amer. Found. for the Blind, 1941. Pp. 215.

HARRIS, R. M. Teachers' Social Knowledge and Its Relation to Pupils' Responses. New York: Teach. Coll., Columbia Univ., 1941. Pp. 89.

CHATTERTON, R. H. Methods of Lesson Observing by Preservice Student-Teachers. New York: Teach. Coll., Columbia Univ., 1941. Pp. 137.

SPADINO, E. J. Writing and Laterality Characteristics of Stuttering Children. New York: Teach. Coll., Columbia Univ., 1941. Pp. 82.

WILLMOTT, J. N. High School Boys Electing Industrial Arts. New York: Teach. Coll., Columbia Univ., 1941. Pp. 71.

TODD, J. E. Social Norms and the Behavior of College Students. New York: Teach. Coll., Columbia Univ., 1941. Pp. 190.

GOLDFARB, W. An Investigation of Reaction Time in Older Adults. New York: Teach. Coll., Columbia Univ., 1941. Pp. 76.

The Journal of Psychology

[illegible]

Carl Murchison

[illegible] the author, the lag in publication should [illegible] four weeks.

(4) [illegible] will receive 200 reprints of his paper, [illegible] the accumulated publications bound [illegible] second half of each volume will contain [illegible] for that volume. Each half volume [illegible] at its stated date.

[illegible] will be charged the actual [illegible] cuts.

(6) [illegible] charges will be as follows:

4	page signature	....	$13
[illegible]	" "	....	24
[illegible]	" "	....	[illegible]
[illegible]	" "	....	[illegible]
[illegible]	" "	....	[illegible]
24	" "	....	[illegible]
28	" "	....	72
32	" "	....	84
[illegible]	" "	....	[illegible]
[illegible]	" "	....	[illegible]
[illegible]	" "	....	[illegible]

[illegible]

(7) [illegible] January, [illegible]

[illegible] volume unit $[illegible]. Complete [illegible] volume plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume Two volumes per year June, 1942
Single numbers $4.00 QUARTERLY Volume 60, Second Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

JUNE, 1942

(Manuscripts are printed in the order of final acceptance)

(over)

Provincetown, Massachusetts
Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass., under the Act of March 3, 1879
Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The Journal of Genetic Psychology, 1942, 60, 197-221.

THE MECHANISM OF VISION: XVII. AUTONOMY OF THE VISUAL CORTEX*

Laboratory of Physiological Psychology, Harvard University

K. S. LASHLEY

The mode of interaction of the different parts of the cerebral cortex is still obscure after a century of experimental study of the brain. The most prevalent conception is that excitations reaching the cortex in sensory projection areas are progressively elaborated as they are transmitted through association fields until they take on a temporal and spacial organization capable of eliciting an appropriate series of adaptive movements when they finally play upon the motor cortex. There are, however, some observations which suggest a certain autonomy in the activities of different cerebral areas. In early studies of cortical function in the discrimination of intensities of light I found that habits of discrimination were retained after destruction of any part of the neopallium except the dorsal convexity of the occipital lobes, a region since identified as the visual projection area (9). Total destruction of this region resulted in loss of the discriminative habits, which could then be reacquired at normal rate. Subsequent destruction of any other part of the cortex did not produce a second loss of the habit (7). Such evidence suggests that the striate areas are the only part of the cerebral cortex involved in the differential reaction to light acquired in the Yerkes' box.

Much the same result for the discrimination of simple geometrical figures was obtained with the "jumping" technique, except that destruction of the striate areas permanently abolished the capacity for detail vision. Lesions in areas surrounding the visual cortex or destruction of the motor areas were without effect upon the learning and retention of habits based upon discrimination of patterns (8, 14). From such studies it seems clear also that no part of the cerebral cortex except the striate areas is essential for learning and retention of simple, visual discriminative habits.

The clinical literature differentiates conditions of post-traumatic amnesia and visual agnosia (psychic blindness) in addition to the

*Received in the Editorial Office on November 14, 1940.

primary visual defects exhibited in scotoma or amblyopia. Nothing clearly comparable to these defects of integration has been discovered in studies of lower mammals. Loss of habits based on discrimination of light after destruction of the striate areas may be comparable to an amnesia (10, 11) but no similar condition has been found for detail vision. Injuries within the striate areas produce retardation in the learning of reactions based upon pattern discrimination, or some inaccuracy of performance, but these defects can be accounted for adequately on the basis of the limitations of the visual field produced by the operation (13). No disturbance of pattern discrimination could be demonstrated as a result of lesions which did not invade the striate areas or optic radiations and, indeed, only involvement of the binocular field with production of a central scotoma produced any significant disturbance in the discrimination of the visual figures used.

The studies have been restricted to discrimination of very simple geometrical forms, lines, or triangles, which are quite readily distinguished and learned by the rat. In human patients visual agnosias may be revealed only in somewhat complex situations and it seemed that for the rat a more difficult problem might reveal transcortical functions which are not evident in the recognition of the distribution of masses of light or the direction of lines (14).

Following this suggestion Kirk (6) attempted to differentiate a postoperative visual amnesia from cortical blindness by the use of a more difficult discriminative problem. He trained rats in a differential reaction between an *F* and its mirror image. He then destroyed part of the occipital cortex of one hemisphere only and tested for postoperative retention. He obtained after operation a partial loss of the habit with a correlation of about 0.50 between the total extent of lesion and the postoperative training required for relearning. The percentages of the striate areas involved in the lesions were estimated by superimposing a diagram of the lesion upon a conventional diagram of the striate area and the values so obtained, representing the actual extent of damage to the visual cortex, were also correlated with the retraining records. The resulting correlation was negligible, amounting to not more than 0.11.

Kirk argued that the restriction of the lesion to one hemisphere ruled out cortical blindness as a cause of the postoperative loss, since at most the lesion could produce only an hemianopia, leaving enough of the visual field intact for recognition of the figures. Such an

assumption is not justified for, in the apparatus used, the rat tends to fixate the edges of the stimulus cards and to react upon the basis of a part of the figure near the inner margin of the frame. Consequently an extensive scotoma might cut off a critical part of the figure and seriously disturb the accuracy of discrimination until a readjustment of fixation was acquired. With symmetrical figures such confusion is especially likely to occur. Limitation of the lesion to one hemisphere does not rule out the possibility that the postoperative disturbances were due to cortical blindness or justify Kirk's conclusion that the loss is evidence for a true amnesia.

Since the total area destroyed correlated significantly with the postoperative scores, whereas the percentage of striate cortex included in the lesion did not, Kirk argued that the cortical fields surrounding the striate areas are important for the formation of the difficult discriminative habit. The method which he used for estimating the extent of damage to visual structures is quite unreliable, since it does not take into consideration either the interruption of the optic radiation or the frequent lateral displacement of remnants of the striate cortex by hernia of the hippocampus. The method gives no real indication of the extent of invasion of the binocular field or of the extent of scotoma produced. The correlation of retraining records with total extent of lesion may have been merely an expression of the probability of greater invasion of the binocular field by large than by small lesions, as I found to be the case in studies of brightness vision (10), and the lack of similar correlation with estimated invasion of the visual field only a result of failure to consider damage to the radiation. The conclusion that the extra-striate cortex is important for the performance of the difficult visual habit is not justified by the data.

The early controversy between Munk (16) and Loeb (15) hinged upon a similar problem of interpretation of experimental results. Munk ascribed the visual loss after destruction of his Area *A* to psychic blindness, an agnosia or disturbance of visual comprehension, whereas Loeb believed that a narrowing of the visual field by scotomas was sufficient to account for the symptoms observed. Crucial evidence upon the question was not presented.

In experiments with animals there is always considerable uncertainty as to what functions a test actually measures. This is especially true in studies of visual defects, since there is no means of accurately mapping scotomas or of controlling visual fixation. To

differentiate amnesia or a semantic defect from effects of direct visual damage, the experiment must fulfill one of two conditions. Either a deterioration of visual function must be demonstrated in cases where there is no damage to the striate cortex or optic path or, if the deterioration occurs in cases with such damage, it must be shown that the animal can meet the sensory requirements of the task without retardation and that the deficiency is in some integrative function which can be distinguished from sensory discrimination.

Accurate determination of damage to the visual system can be made by analysis of retrograde degeneration in the lateral geniculate nucleus (9). No other method seems reliable. Even careful mapping of the remnants of the striate cortex by cytoarchitectural criteria may give misleading results. I have cases in which inflammatory changes have so obscured the cellular picture that no trace of the striate cortex can be made out and which nevertheless retained some detail vision, correlating with islands of normal cells in the lateral geniculate nucleus. Anatomic studies have not revealed any direct visual connections to the cortex except by way of the geniculo-striate system, so a reconstruction of the nucleus gives a fairly trustworthy index of the degree of interference with primary visual functions.[1] For cases with direct injury to the visual system it is impossible, with available methods, to distinguish between effects of scotoma and disturbance of gnostic functions, except where the latter are based upon discriminative reactions which are themselves undisturbed by the lesions.

The experiment reported below was designed to test the effect of injuries to the neopallium upon a visual function which seems to involve some degree of integration beyond that required for discrimination of visual figures. The function is that which I have called a "conditional reaction" (12). It requires a choice of one or the other of two figures according to the character of the background upon which the figures are shown. Additional tests were included to determine whether or not the animals were capable of discriminating the figures which formed the basis for the conditional reaction.

Methods

Pigmented rats of a strain derived from a cross of local albino with trapped gray animals were used. They were quite wild and

[1]For a description of the visual paths and cortex in the rat see (9, 18).

required a week to 10 days of preliminary handling to adapt them to the apparatus. Microphakia occurs in the strain, so it was necessary to remove and examine the lenses of all animals used in the experiment. Only one animal with this defect, No. 10, was found. About equal numbers of males and females were used, ranging in age from 130 to 160 days at the beginning of training.

Operations were performed under ether anesthesia. Areas of the cortex were destroyed by thermocautery or the fiber connections between different areas were cut with a small curved knife. In the series of cases the greater part of the neopallium was explored, either by direct destruction or by interruption of possible connections with the area striata. In general I attempted to avoid direct damage to the cortical area for the binocular field, but this area or its radiation was invaded in a number of cases.

At the termination of the experiment the animals were sacrificed and serial sections through the lesions and thalamus stained with thionin. The lesions were reconstructed by the usual graphic method. Degeneration in the lateral geniculate nuclei was charted. The distribution of degeneration in each nucleus at midsection is shown diagrammatically along with the charts of cerebral lesions in Plates 1 and 2.

An enclosed jumping stand, painted black, was used for training. The animals were required to correct their errors, that is, in case of a jump to the negative stimulus the cards were left in the same position and the animal made to jump again until he chose the positive card. If he persisted in jumping to the negative card he was constrained and forced to jump to the positive after 10 such "repetitive errors." In the scores a "trial" represents a jump to the positive stimulus, together with any preceding errors. In the tables the score of "errors" actually represents the number of trials in which one or more false jumps were made. This method is used for more ready comparison of scores with expectation from chance.

Training was continued in Tests 1, 2, 3, and 5 with 10 trials per day to a criterion of 20 consecutive errorless trials or to 300 trials, in case the criterion was not met earlier. In a few cases where the rat was making better than chance scores at 300 trials training was prolonged beyond this amount.

1. *Sequence of Tests*

The tests used are listed below in the order in which they were

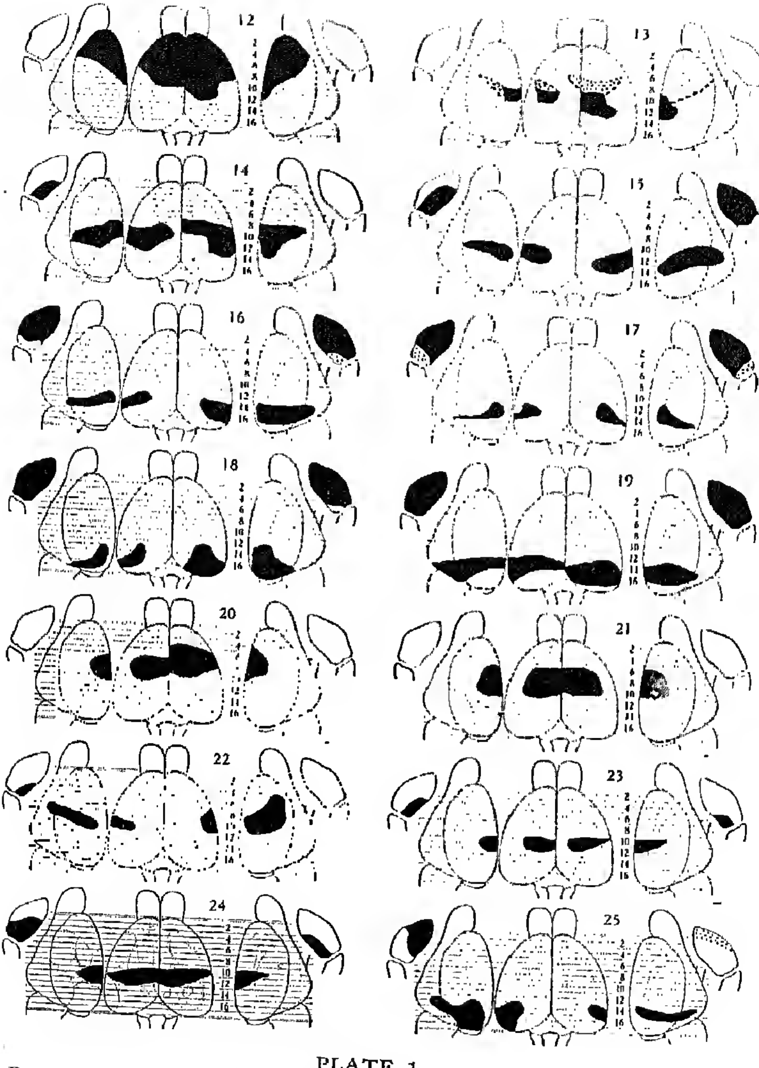

PLATE 1

DIAGRAMS SHOWING THE EXTENT OF INJURY TO THE CORTEX AND VISUAL SYSTEM IN INDIVIDUAL CASES

Extent of cortical destruction is shown in solid black. In Figure 13 stippled areas represent deep cuts through the corpora striata. Small diagrams show the extent of degeneration (black) in a mid-section through each lateral geniculate nucleus. Coarse stippling represents partially degenerated areas in which a few normal cells remain. Figures 12 to 19, cases which failed Test 2; Figures 20 to 27, cases which failed Test 5; Figures 28 to 39, cases which passed all tests.

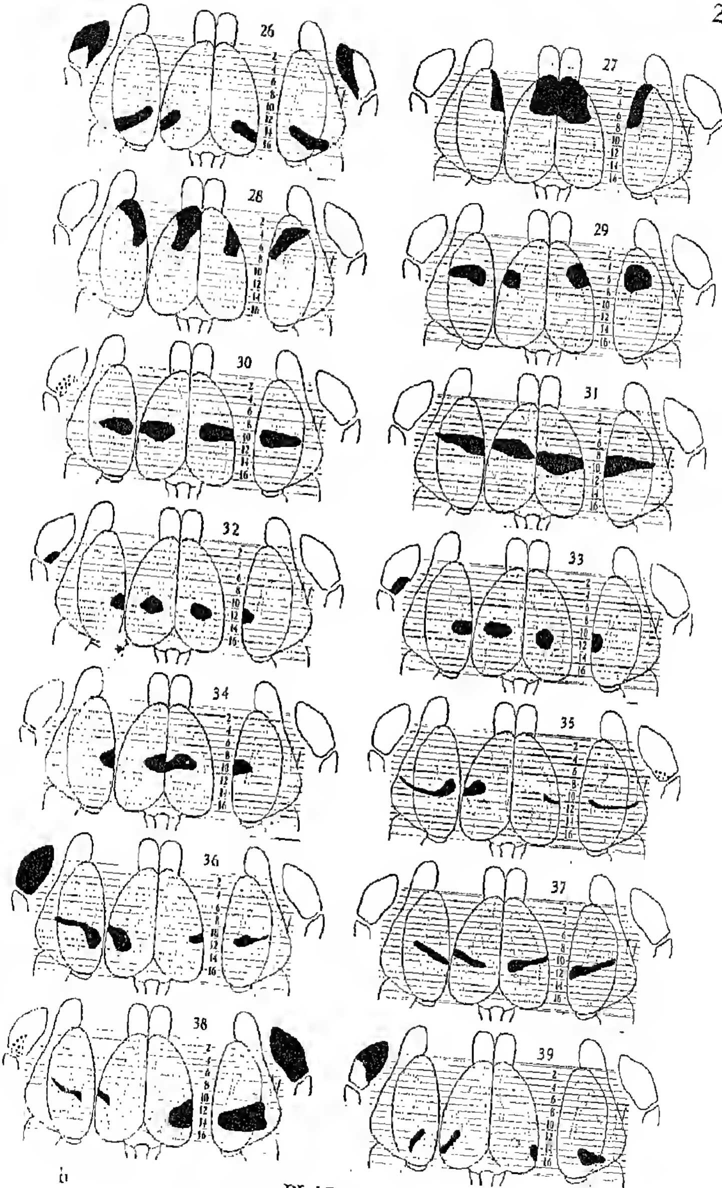

PLATE 2

given, following training to jump to the open doors of the apparatus.

1. To choose a 10 cm. white square on a black card and avoid a black card (Figure 1, *a*). This is a very easy problem and is frequently learned after a single error by normal animals.

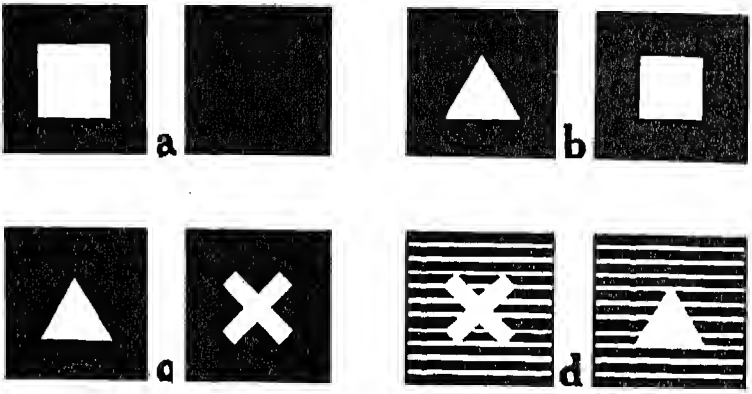

FIGURE 1

PATTERNS USED IN TRAINING

a, Test 1; *b*, Test 2; *c*, Test 3; *d*, Tests 4 and 5; *c* and *d* alternately, Test 6. Cards 15 cm. square.

2. To choose an equilateral triangle of 112 sq. cm. area and to avoid a square of equal area, white figures on black ground (Figure 1, *b*); a very difficult discrimination, probably because of the similarity of the base-lines of the figures. This and the preceding test provide a measure of the influence of the cerebral lesion upon discriminative processes.

3. To choose the same triangle and avoid an *X* of equal area (Figure 1 *c*). This substitution was made to provide a greater contrast between the figures to be used for the conditional reaction. The transfer from Test 2 was made readily by most animals.

4. Critical trials, a test for preference, between the same triangle and *X* when presented on a striated ground (Figure 1, *d*). Twenty trials were given, with reward for every jump, irrespective of the figure chosen. The records are expressed as the percentage of trials in which the triangle was chosen. A preference for one figure indicates an identification of the figures on the striated ground with those

learned on black ground and also visual capacity sufficient for discrimination under the new conditions.

5. To choose the *X* and avoid the triangle on striated ground (Figure 1, *d*). This test involves reversal of the initial direction of training and should reveal a reduction of plasticity, if it occurs in operated animals. Errors in the first 20 trials of the test are recorded separately, since they sometimes show transfer of the effects of previous training not revealed by the critical trials.

6. Alternate training to 10 consecutive errorless trials with the figures on black ground, triangle positive, and on striations, *X* positive, until spontaneous reversal of reaction occurred with each ground. The criterion of spontaneous reversal was five errorless trials immediately following the change in ground and not more than one error in the succeeding five trials. The scores for this test record the number of reversals of training required before the criterion of spontaneous reversal was met for both grounds successively. Thus a score of 5 means that the animal had been given the initial training with *X* positive on striations (1st reversal), then with triangle positive, black ground (2nd reversal), *X* positive on striations (3rd reversal), triangle positive on black (4th reversal), *X* positive on striations (5th reversal), and had then chosen the triangle on black ground without error on the 6th reversal and the *X* on striations without error on the 7th. In addition to the number of reversals of training, the total number of errors made after the first reversal (initial training with striations) is recorded.

2. *Incidental Observations*

Detailed notes were kept upon individual differences in behavior. Evidences of sensorimotor disturbance, timidity, reluctance or refusal to jump, characteristic postures and the like, were recorded. Sensorimotor difficulties, indicated by frequent falls and misplacing the feet in preparation for jumping, correlate with lesions in the electro-stimulable areas and with refusal to jump after repeated errors (v. i.). Other individual peculiarities did not appear to be significant.

Experimental Results

The records for the tests are summarized in Tables 1 to 5. All animals without exception learned to jump and to choose the white square opposed to the black card (Test 1). On the basis of subsequent records they have been divided into the following groups

TABLE 1

TEST SCORES OF NORMAL ANIMALS

For description of the tests see page 204. *Tr.*, trials and *E.*, errors preceding the criterion of 20 successive errorless trials; *% crit.*, percentage of 20 critical trials in which the triangle was chosen; *% error*, percentage of errors in first 20 trials in first reversal of training; *rev.*, number of reversals of training required before spontaneous reversal occurred; *B.*, animals which balked in training; *F.*, animals which failed. The last column is the number of errors in reversed training after Test 5.

No. of test	1		2		3		4	5			6	
	Tr.	*E.*	*Tr.*	*E.*	*Tr.*	*E.*	*% crit.*	*Tr.*	*E.*	*% error*	*Rev.*	*E.*
Rat No.												
1	27	9	136	60	6	3	50	226	99	60	4	14
2	61	15	123	54	0	0	60	163	67	70	3	29
3	2	2	115	49	9	5	80	136	68	80	3	58
4	2	2	148	58	18	4	70	203	111	80	3	34
5	45	9	110	42	33	12	60	208	99	100	4	65
6	12	2	173	76	38	14	50	117	70	100	7	44
7	47	14	148	54	37	8	70	230	79	50	4	15
8	31	11	108	46	0	0	50	143	54	40	4	15
9	29	9	134	58	7	3	50	191	79	90	3	39
10	25	3	159	53	47	14	50	390	157 *F*	60	–	—
11	49	9	238	105	90	26	80	300	172 *F*	60	–	—

TABLE 2

SCORES OF OPERATED ANIMALS WHICH FAILED TEST 2

(Abbreviations as in Table 1.)

No. of test	1		2	
	Tr.	*E.*	*Tr.*	*E.*
Rat No.				
12	10	1	—	— *B*
13	6	1	—	— *B*
14	28	2	40	23 *B*
15	10	4	300	152 *F*
16	60	18	300	159 *F*
17	56	18	300	167 *F*
18	0	0	300	159 *F*
19	55	14	300	168 *F*

which will be discussed separately; normal animals, operated animals which failed in Test 2, operated animals which failed in Test 5, and operated animals which passed all tests.

1. *Normal Animals*

(Table 1). Two normal rats failed in Test 5. One of these, number 10, had a severe microphakia with lenses of less than half normal diameter. The other, number 11, had normal lenses but did so badly in all tests of detail vision as to suggest some other visual defect. The scores of these two cases are therefore omitted in computing averages for the group. The average scores for the others are given in Table 5 in comparison with the scores for the operated groups.

2. *Operated Cases which Failed in Discrimination of Triangle and Square*

Eight cases are included in this group. Their records are given in Table 2 and average scores in Table 5. Diagrams of the cerebral lesions are presented in Plate 1, Figures 12 to 19. All of these animals learned black-white discrimination with an average of 28 trials and 7 errors, a score almost exactly equal to that of the normal group. In Test 2 their behavior divided them sharply into two groups. Nos. 12, 13, and 14 refused to jump, the first two within 10 trials, No. 14 after 40 trials. They would still jump readily to a single figure but could not be forced when the two figures appeared. These three animals had extensive lesions within the sensorimotor fields, without significant visual involvement. Their behavior indi-

TABLE 3

SCORES OF OPERATED ANIMALS WHICH FAILED IN TEST 5
(Abbreviations as in Table 1.)

No. of test	1		2		3		4	5		
	Tr.	*E.*	*Tr.*	*E.*	*Tr.*	*E.*	*% crit.*	*Tr.*	*E.*	*% error*
Rat No.										
20	18	3	218	87	0	0	70	110	50 *B*	45
21	4	2	60	21	14	2	50	300	102 *BF*	50
22	20	2	200	78	39	9	80	300	164 *F*	90
23	6	3	109	61	77	32	70	300	156 *F*	60
24	0	0	205	71	5	1	85	300	147 *F*	60
25	12	4	273	82	0	0	80	300	132 *F*	70
26	25	8	150	77	160	70	80	300	163 *F*	75
27	0	0	130	50	0	0	B	—	—	—

TABLE 4
SCORES OF OPERATED ANIMALS WHICH PASSED ALL TESTS
(Abbreviations as in Table 1.)

No. of test	1		2		3		4	5			6	
	Tr.	*E.*	*Tr.*	*E.*	*Tr.*	*E.*	*% crit.*	*Tr.*	*E.*	*% error*	*Rev.*	*E.*
Rat No.												
28	0	0	111	39	0	0	70	135	61	70	4	25
29	14	8	101	42	12	4	50	147	45	70	4	14
30	28	6	159	74	26	6	70	190	81	50	2	27
31	48	21	99	48	19	5	50	230	115	100	5	57
32	18	5	159	44	13	3	40	80	44	80	5	11
33	0	0	153	62	16	4	50	77	24	60	3	3
34	20	2	107	28	8	2	60	51	17	40	1	0
35	7	2	104	49	67	10	70	176	63	60	6	33
36	9	2	250	112	83	23	40	336	119	50	2	17
37	25	8	168	74	14	4	60	238	67	50	7	49
38	50	15	262	113	0	0	50	152	72	80	7	63
39	0	0	170	52	0	0	40	162	55	60	7	22

TABLE 5

COMPARISON OF THE AVERAGE SCORES OF THE DIFFERENT GROUPS

(Abbreviations as in Table 1.)

No. of test	1		2		3		5		6	
	Tr.	*E.*	*Tr.*	*E.*	*Tr.*	*E.*	*Tr.*	*E.*	*Rev.*	*E.*
Normal	28.4	8.1	132.8	55.2	16.5	5.5	179.7	80.7	3.9	34.8
Operated										
Failed Test 2	28.1	7.3								
Failed Test 5	10.6	2.7	156.2	65.9	36.8	14.2				
Passed all	18.2	5.8	153.6	61.4	21.5	5.0	164.5	63.6	4.4	26.8

cated that they could distinguish visually between the stimuli of Tests 1 and 2, that is, the presence of one or of two figures.

The other animals of this group jumped readily for 300 trials in Test 2, but their records did not improve above chance. In all of them there was bilateral destruction of the binocular field with extensive damage in the temporal visual fields as well, amounting in three of the cases to practically total destruction of both optic radiations. The inability of the animals with such lesions to distinguish visual patterns, although retaining discrimination of black and white, is in accord with the results of previous experiments.

3. *Operated Cases which Failed in Discrimination of X and Triangle on Striated Ground (Test 5)*

The records of the animals in this group are given in Table 3 and diagrams of the lesions in Plates 1 and 2, Figures 20 to 27. No. 27, after errorless transfer from triangle vs. square to triangle vs. *X*, suddenly began to jump against the left restraining wall of the apparatus and could not again be induced to jump to the stimulus cards. Sections of the brain showed recent hemorrhage and inflammatory processes in the head of the right caudate nucleus. Two of the remaining seven animals balked during training in Test 5. No. 20 became progressively more hesitant and after 110 trials could not be forced to jump again. No. 21 balked after the first 10 trials and could not be forced to jump by blows, electric shocks, or air blasts. She would jump, however, if pushed over the edge of the platform, and was carried through 300 trials by this method. She occasionally made better than chance scores but clearly on the basis of cues from the experimenter. These two animals both had extensive lesions in the sensorimotor areas without damage to the visual system.

No. 22, 23, 24, 25, and 26 all jumped readily but were still making chance scores after 300 trials. All of them, except No. 22, had bilateral involvement of the binocular field, but less extensive than appeared in those animals which failed in discrimination of square and triangle.

No. 22 had only a slight invasion of the left optic radiation, sufficient to produce a small unilateral scotoma. The cortical lesions are almost identical with those in No. 29 (v. i.), which readily passed all tests, and seem insufficient to account for the animal's failure.

4. *Operated Cases which Passed All Tests*

Twelve animals with brain lesions passed all tests. Their records are presented in Table 4 and in Plate 2, Figures 28 to 39. None of the animals of this group was significantly inferior to the normal controls in all tests. The lesions cover the principal areas of the cortex, or separate them from the striate areas.

INTERPRETATION OF EXPERIMENTAL DATA

Save for one animal, No. 22, the experiment gives a consistent picture. The lesions in this one exceptional case, which failed Test 5, are almost identical with those in No. 29, which passed all tests easily. The failure of No. 22 can therefore scarcely be ascribed to the lesions and, since one of the 10 normal controls also failed Test 5, it seems justifiable to disregard No. 22.

1. *Rats that Refused to Jump*

Six animals balked at some stage in the training. These were Nos. 12, 13, 14, 20, 21, and 24. No. 24 balked in the early training on Test 2, but jumped spontaneously again when returned to Test 1 and thereafter jumped readily in all tests. The lesions in these animals in every case invaded the electrostimulable area for the limbs and centered on the area described by Brooks (1) as essential for hopping and placing reactions. Similar symptoms were recorded for all six cases during the first days of training; frequent falls from the platform, misplacing the feet, a fine tremor when getting set to jump, and long hesitation. All jumped spontaneously, however, to the open doors and to the single white square of Test 1. The motor symptoms improved rapidly, to leave only a slight clumsiness suggestive of a sensory defect.

The behavior of No. 21 is fairly typical of the group. She jumped readily in Tests 1, 2, 3, and 4 and her learning scores were better than those of any normal animal. After the first 10 trials of Test 5 she refused to jump and thereafter jumped to the striated figures only when pushed off the platform. However, she would jump readily to the open doors, or to the white square vs. black and, if one of the striated cards were displaced slightly, so as to show a slit 5 mm. in width between it and the frame, she would jump to it quickly and accurately. She also jumped again, with some urging, to triangle vs. square. The lesion covered the electrostimulable area for the limbs.

It is difficult to evaluate failure in the tests in terms of such behavior. There is no suggestion of a primary visual defect. Sensori-motor defects did not interfere with learning of the easier tasks. The animals showed no retardation in the problems which they learned. They simply quit after a number of failures which are insufficient to discourage a normal animal. The behavior suggests that of some human patients with cerebral lesions who appear normal in situations which they can solve easily but seem unable to make the necessary effort to achieve a difficult task. The behavior does not suggest any defect specific for visual tasks and can scarcely be taken as evidence of a visual agnosia.

2. *Failure Resulting from Primary Visual Defects*

The failure of all rats, except the seven discussed above, can be accounted for in terms of primary visual defects. Those which failed discrimination of triangle vs. square (Nos. 15 to 19) had practically complete destruction of the optic radiations. Those which passed this test but failed Test 5 (Nos. 23 to 26) had extensive bilateral damage so located as to produce an extensive central scotoma. The effects were essentially the same when the injury was to the optic radiation at some distance from the striate cortex as when the striate cortex itself was involved. No animal which passed all of the tests had such bilateral damage to the primary visual paths. Failure in the tests is thus clearly a function of a primary defect, a cortical blindness involving parts of the fields of both temporal retinae. Whether or not there may have been an additional integrative difficulty is not revealed by the data. The normal performance of the operated animals which passed the tests and in which many of the lesions were similar except that there was not *bilateral* involvement of the radiations argues against any condition resembling agnosia.

3. *Operated Rats that Passed the Tests*

a. Normal performance in tests passed. There is little evidence of subnormality in the scores for any test in those animals which were able to pass the test at all. The scores, in general, show an all or nothing function; either the animal made no improvement during 300 trials of training or he made a record well within the range of the normal controls. The average scores of normal and operated groups are shown in Table 5. In acquisition of the reaction

to white vs. black the averages of all operated groups are superior to those of the normal controls. This superiority is probably an expression of the taming incident to postoperative care and handling. In Tests 2 and 3 the operated animals are slightly inferior to the normal, but when the individual records are examined it appears that the poorer scores are due chiefly to the animals with extensive scotomas. In Test 5 the difference is in favor of the operated group to about the same amount. In Test 6, the development of the conditional reaction, the normal group required 11 per cent fewer reversals of training but made 30 per cent more errors than did the operated group.

The variability of the operated groups exceeds that of the controls. In every test the best scores, as well as the worst, were made by animals with cerebral lesions. The variability of the operated cases seems to be the result of three factors: taming by post-operative nursing and a somewhat greater caution resulting from physical handicap make for lower scores, whereas scotomas reduce efficiency. Those operated cases which made consistently bad records practically all had severe primary visual defects. The differences in variability of operates and normals are, however, neither very great nor statistically reliable.

b. The conditional reaction. All the animals which learned the first reversal of reaction (Test 5) came through to spontaneous reversal according to the character of the ground with little additional practice. The average number of reversals required by the normal animals was 3.9, that for the operated cases was 4.4. Four operated animals did as well as the best score of the normal ones. The operated animals made significantly fewer errors before reaching the criterion (normals, 35; operated, 27). The operated animals were less disturbed in the formation of the conditional reaction than in the first difficult problem of discrimination.

The failure of the cerebral lesions to interfere with the formation of the conditional reaction in any animals which passed Test 5 may be interpreted in any of three ways: (*a*) The conditional reaction may be merely the formation of two independent habits, not involving any processes beyond simple discrimination. (*b*) Animals with brain lesions may be less able than normal ones to identify the figures in the original with those in the reversed situations and thus their real inferiority in processes of generalization may be obscured by the fact that they have less negative transfer to overcome. (*c*)

Complex and difficult visual functions may be no more dependent upon the extra-striate cortex than is simple discrimination.

1. As in an earlier study of the conditional reaction (12) it seems certain that the training in Test 5 involved a true reversal of the habit of Test 3. On the basis of the critical trials (Test 4) and of the first 20 trials of reversed training seven of nine normal animals and seven of the 12 operated showed a preference for the triangle when triangle and *X* were presented on striated ground. On the next reversal (black ground, triangle positive after training with *X* positive) only one animal, No. 34, chose the triangle without initial error. In the normal group there was also a significant retardation in the learning of Test 5 over the practice required in Test 2, an indication of negative transfer from Test 3. Thus the majority of animals gave evidence for an identification of the figures on black ground with those on striations and for an interference of the two habits. For most animals the conditional reaction does require something more than the formation of two independent associations.

2. Only 33 per cent of normal and 25 per cent of operated animals, which passed all tests, chose the triangle when first presented on striated ground in 70 per cent or more of 20 critical trials. In the first reversed training with punishment for choice of the triangle, 66 per cent of the normal group and 42 per cent of the operated showed a definite preference for the triangle during the first 20 trials of training.[2] Combining the results of the two tests, 77 per cent of normal and 58 per cent of operated cases gave evidence of negative transfer. Thus a smaller proportion of operated than of normal animals seemed to identify the figures in the two situations. In the operated group (Table 4) the average scores for the seven animals which showed initial preference for the triangle were 4.7 reversals and 32 errors. For the five animals which did not show such preference the scores were 4.0 reversals and 18.2 errors. The

[2]These scores are of interest for evaluation of the method of critical trials, in which a preference for one of a pair of stimuli is tested by allowing the animal to choose between them with reward for every choice. After training to choose triangle vs. *X*, Nos. 5, 6, and 31 made chance scores when tested with the same figures on striated ground. When, however, reversed training was begun, with punishment for choice of the triangle, these animals made 100 per cent error for the first 20 trials. A chance score in critical trials thus cannot be accepted as evidence that the new stimuli are not identified with others on which the animal has been trained.

influence of negative transfer upon the number of reversals is slight but the error score is significantly affected.

The relative absence of negative transfer among the animals of the operated group is, however, only sufficient to account for the superiority of the group over the normal controls. The scores of the operated cases which did show negative transfer are 4.7 trials and 32 errors. The scores of the normal group are 3.9 reversals with 35 errors. Thus the operated animals which did identify the figures in the initial and reversed situations were not significantly inferior to the normal controls and the failure of some animals of the operated group to make a negative transfer did not obscure a general inferiority in capacity to acquire the conditional reaction.

The evidence from the training records shows that the only animals in the groups with cerebral lesions which were inferior to the normal controls were those in which the defects can be referred either to a form of sensorimotor disorder, associated with a reluctance to jump in situations requiring difficult discrimination, or to a primary visual deficiency. The reluctance to jump is almost certainly a result of motor disorder only. I have observed a similar balking after a few falls in animals with action tremor from cerebellar lesions and in these the symptom can scarcely be ascribed to any disorder of visual or gnostic functions. The failure or retardation of animals with scotomas, as inferred from degeneration in the lateral geniculate nuclei, is proportional to the extent and central position of the scotoma, irrespective of whether the lesion invades the striate cortex or interrupts the radiation in the internal capsule. There is therefore no reason to assume that failure in the tests has involved any semantic defect in addition to the cortical blindness.

4. *Anatomic Considerations*

The studies of Clark (3, 4, 5) and of Waller (18) have outlined roughly the sensory projection areas of the rat's cortex. The anterior group of nuclei, probably olfactory, is projected to the dorsomedial surface of the hemisphere, probably to the infraradial and retrosplenial regions of Rose (17). The ventral group of nuclei, relays for cutaneous and somatic sensory functions, project to the latero-dorsal and lateral regions of the cortex, a field corresponding roughly to the postcentral and temporal regions of Rose. These regions, as well as the electrostimulable cortex, were severed from

any direct connection with the striate cortex in several of the operated rats which passed all tests (Nos. 30, 31, 34, 35, and 36).[3]

The lateral nucleus of the thalamus is believed to be concerned in integration of thalamic functions. Its cortical connections are with the parietal area. In the rat the nucleus is divided into anterior and posterior parts. The field of the anterior part is nearly coextensive with the stimulable field for the limbs. Direct connections between this field and the striate area were more or less completely interrupted in Nos. 32, 33, and 34. In No. 32 there was nearly total bilateral degeneration of both lateral nuclei, pars posterior. Connections with the supposedly associative parietal field are therefore not essential for the visual functions studied. In No. 39 the auditory cortex was extensively damaged on the right together with nearly complete interruption of the left optic radiation. The animal's score was normal in all tests. There remains unexplored a limited area along the lateral margin of the striate cortex which Cajal (2) has described as lacking projection fibers. It perhaps corresponds to Rose's Area occipitalis, which lies above the optic radiation and cannot be extensively damaged without interrupting the radiation and producing a cortical blindness.

Except for the unexplored part of this area occipitalis, the sensory, motor, and associative fields of the neopallium have all been severed, in one or another animal, from direct connection with the striate cortex.

5. *Autonomy of the Visual System*

The experiment thus shows that animals with interruption of direct connections between the striate areas and other parts of the neopallium may acquire difficult visual associations at a normal rate. The only demonstrable interferences with the functions were clearly the result of motor disturbance or of extensive central scotomas. If the functions studied do involve interaction between the visual cortex and other cortical areas, then this interaction must be either by way of diffuse connections through whatever bridges of cortex remain intact or through indirect connections to and from lower centers.

[3]It is doubtful whether there are long transcortical association fibers in the rat. After lesions in the striate area Marchi preparations show the optic radiation and a bundle of fibers passing through the splenium to the opposite hemisphere but no other degenerated fibers can be followed for more than 0.5 mm. beyond the edge of the lesion.

satisfactory substitutes for these concepts have been proposed. It seems certain from many lines of evidence that the organization of visual impressions in perceptual patterns, illustrated by figure-ground relations or the identification of form, is a primitive function which is disturbed only by direct damage to the visual cortex. The conditional reaction, although it requires an additional step in generalization beyond discrimination of form, is no more dependent upon transcortical activities than is the latter function. In what respects, qualitative or quantitative, the visual functions which are disrupted in man by extrastriate lesions differ from those which have been studied in the rat remains to be determined.

Summary

After a variety of cerebral lesions rats were trained, by a technique requiring jumping against the stimulus cards, in a visual reaction based upon a difficult discrimination and also in a problem which calls for a spontaneous reversal of reaction to a pair of figures in accordance with the background upon which they are displayed. Some rats which jumped readily with an easy problem refused to jump when discrimination became difficult. Such behavior was associated with lesions centering on the electrostimulable area for the fore and hind legs. In all other animals which failed in the tests there was anatomic evidence of extensive bilateral central scotomas.

The operated animals which passed the tests were not significantly inferior to a group of normal controls. In general, either the animal made no improvement during 300 trials of training or his learning score fell within the normal range. No rat which passed all of the tests was consistently retarded in all, and those which were seriously retarded in some tests had extensive unilateral scotomas.

In the series of cases which passed all tests various direct connections between the visual cortex and other parts of the neopallium were interrupted, and in one or another case all the major cytoarchitectural fields except the striate areas were seriously damaged, without deterioration of the functions.

From these facts it is concluded that the functions of the visual cortex in the performance of difficult visual discriminations do not depend upon any direct or specific transcortical connections with other regions of the neopallium. If interaction between the striate cortex and other parts of the neopallium occurs, it must be either by way of connections through subcortical regions or by diffuse con-

duction through any cortical field remaining intact. Evidence from previous studies of easier discriminations makes it probable that the normal activities of the visual system in visual functions are not dependent upon interaction with other parts of the neopallium.

References

1. Brooks, C. M. Studies on the cerebral cortex: II. Localized representation of hopping and placing reactions in the rat. *Amer. J. Physiol.*, 1933, **105**, 162-171.
2. Cajal, S. Ramón Y. Histologie du système nerveux de l'homme et des vertébrés. Paris: 1911.
3. Clark, W. E. Le Gros. An experimental study of thalamic connections in the rat. *Philos. Trans. Roy. Soc. Lond.*, 1932, **222**, B, 1-34.
4. Clark, W. E. Le Gros, & Boggon, R. H. On the connections of the anterior nucleus of the thalamus. *J. Anat.*, 1933, **67**, Pt. II, 215-226.
5. ———. On the connections of the medial cell groups of the thalamus. *Brain*, 1933, **56**, 83-98.
6. Kirk, S. A. Extra-striate functions in the discrimination of complex visual patterns. *J. Comp. Psychol.*, 1936, **21**, 145-159.
7. Lashley, K. S. Studies of cerebral function in learning: IV. Vicarious function after destruction of the visual areas. *Amer. J. Physiol.*, 1922, **59**, 44-71.
8. ———. The mechanism of vision: IV. The cerebral areas necessary for pattern vision in the rat. *J. Comp. Neurol.*, 1931, **53**, 419-478.
9. ———. The mechanism of vision: VIII. The projection of the retina upon the cerebral cortex of the rat. *J. Comp. Neurol.*, 1934, **60**, 57-79.
10. ———. The mechanism of vision: XII. Nervous structures concerned in the acquisition and retention of habits based on reactions to light. *Comp. Psychol., Monog.*, 1935, **2**, No. 52, 43-79.
11. ———. The mechanism of vision: XIII. Cerebral function in discrimination of brightness when detail vision is controlled. *J. Comp. Neurol.*, 1937, **66**, 471-480.
12. ———. Conditional reactions in the rat. *J. of Psychol.*, 1938, **6**, 311-324.
13. ———. The mechanism of vision: XVI. The functioning of small remnants of the visual cortex. *J. Comp. Neurol.*, 1939, **70**, 45-67.
14. Lashley, K. S., & Frank, M. The mechanism of vision: X. Postoperative disturbances of habits based on detail vision in the rat after lesions in the cerebral visual areas. *J. Comp. Psychol.*, 1934, **17**, 355-380.
15. Loeb, J. Die Sehstörungen nach Verletzungen der Grosshirnrinde. *Arch. f. d. ges. Physiol.*, 1884, **34**, 67-172.
16. Munk, H. Über die Funktionen der Grosshirnrinde. Berlin: 1881.

17. ROSE, M. Cytoarchitektonischer Atlas der Grosshirnrinde der Maus. *J. f. Psychol. u. Neurol.*, 1929, **40**, 1-51.

18. WALLER, W. H. Topographical relations of cortical lesions to thalamic nuclei in the albino rat. *J. Comp. Neurol.*, 1934, **60**, 237-269.

Yerkes Laboratories of Primate Biology
Orange Park, Florida

The Journal of Genetic Psychology, 1942, 60, 223-257

VISUAL BRIGHTNESS DISCRIMINATION IN THE CAT AS A FUNCTION OF ILLUMINATION*[1]

Department of Psychology, Tufts College

LEONARD C. MEAD

A. INTRODUCTION

The purpose of the present investigation is to secure a complete curve of visual intensity discrimination as a function of illumination in a typical laboratory mammal, the cat. Prior studies of intensity discrimination throughout the brightness range of the eye have been confined to experiments upon the human individual and upon invertebrates. Previous experiments made upon infra-human vertebrates have secured limens usually at only one brightness level, or, at most, over a small range of illumination. Accordingly, precise data are lacking concerning this visual capacity in infra-human vertebrates, the knowledge of which may be of significance in formulating ideas pertaining to the operation of the retina and nervous system in vision.

One of the most descriptive theoretical systems pertaining to intensity discrimination has been formulated by Hecht (18, 19, 20, 21, 22). His theoretical photochemical system has been found to apply to the data obtained with various invertebrate organisms and with man, and is suggestive of the part played by peripheral mechanisms in mediating this visual function. The present paper will attempt to test the generality of Hecht's theoretical formulations by applying his system to data obtained with the cat.

Besides providing important quantitative information relative to

*Accepted for publication by Leonard Carmichael of the Editorial Board, and received in the Editorial Office on November 14, 1940.

[1]The main portion of the experimental work reported in this paper was performed in the Laboratory of Psychology at the University of Rochester, and was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology in the Graduate School of the University of Rochester. Grateful acknowledgment is made to Dr. Karl U. Smith for assistance and guidance in the completion of this work. Dr. Leonard Carmichael kindly read over the manuscript and offered several valuable suggestions.

the psychological and physiological nature of vision, the study of intensity discrimination also offers a method for the investigation of the functions of the central mechanisms related to this sensory field. The use of the extirpation method as a means of clarifying problems of the neural basis of sensory functions necessitates at the outset the establishment of quantitative descriptions of these functions in the normal animal. The present paper may therefore also be considered as a preliminary in the investigation of the rôle of the central neural centers mediating differential brightness vision.

B. Intensity Discrimination in Phylogeny

1. *Invertebrates*

The literature dealing with the effects of changes in light intensity on the behavior of invertebrate organisms is extensive, although most of the observations are more qualitative than quantitative. Only in comparatively recent years have precise threshold measurements been obtained on certain of these lower forms, and the present discussion will be confined to some of these more quantitative investigations.

Differential responses on the part of various invertebrate organisms have served as the objective criteria by which the effects of light of different intensities on behavior can be measured. Thus Folger (9) measured the reaction time and latent period of the response of amoeba to light varying from 3 to 16 candle power in intensity. The reaction time was found to be inversely proportional to the stimulating intensity, whereas the latent period increased at low brightnesses and then decreased as the brightness was raised.

The photochemical law, which states that a constant amount of energy is required to produce a given photolytic effect regardless of its distribution in time (Bunsen-Roscoe law), has been found to hold for the hydroid Eudendrium (Loeb and Ewald, 30). The freshly regenerated polyps were exposed to a 40 candle power lamp placed at different distances from the aquarium. The time required to produce equal effects (bending of half the colony toward the light) was determined at each distance, and the light intensity calculated by the inverse square law. Hecht (13, 14, 15, 17), Folger (10), and Pieron (43) have all demonstrated that the Bunsen-Roscoe law holds within limits for the clam, *Mya arenaria.*

The reaction time of the jelly-fish, *Medusa Goneonemus,* was found to vary inversely with light intensity. Yerkes (62) noted that the swimming animal, when exposed to weak daylight responded in 9.4 seconds, to strong daylight in 7.0 seconds, and to strong sunlight in 5.5 seconds.

Although the above studies offer direct measurements concerning reactions to intensity changes in certain invertebrate forms, they have not been expressly concerned with the capacity of these animals to discriminate intensity differences as a function of illumination. The following experiments, which have used the tropistic responses of these animals to light, are directly related to this problem.

Crozier and Cole (8) have measured the photic orientation of the slug Limax as it creeps geotropically along an inclined plane. The angular deflection from a line bisecting the inclined plane vertically was found to be proportional to the logarithm of the stimulating intensity. The circus movement effect, expressed as the amount of turning per unit length of path, was found to follow Weber's law[2] from 0.0 to 3.0 meter candles.

Hecht and his collaborators (17, 18, 19, 20, 21, 22, 23, 56, 59, 60, 61) have contributed notably to the problem of visual intensity discrimination, utilizing the data derived from studies on certain lower organisms and man as a means of testing the validity of his theoretical assumptions regarding the photochemical nature of brightness discrimination. In an early study with the clam, Hecht

[2]The use of the Weber fraction in the present and remaining portions of this paper should not be understood as implying the validity of the Weber-Fechner "law", i.e., that $\frac{\Delta I}{I}$ is a rectilinear function of log I. This topic has been adequately discussed by Hecht (16) for vision. It is generally agreed that the fraction varies continuously but probably not rectilinearly with log I. Thus, as Hecht (17) has correctly stated, the theoretical formulation of $\frac{\Delta I}{I}$ as a constant function of illumination is false, and the problem that now remains is how this ratio *varies* with intensity. As here used, the expression $\frac{\Delta I}{I}$ is, of course, equal to the difference between the standard and variable stimulus divided by the standard, i.e.,

$$\frac{I - I_1}{I_1} = \frac{\Delta I}{I}$$

(13, 14, 15) adapted the animal to a certain intensity of light and then determined the increase in intensity required to bring about siphon retraction. The amount of increase necessary to produce this response in a given time varies with the adaptation level, and the ratio between these intensities can be expressed in the form $\frac{\Delta I}{I}$. Plotting this fraction against the adapting intensity yields a curve of intensity discrimination.

Hecht and Wald (23) measured visual intensity discrimination in Drosophila using a method in which the observer noted the increase in illumination of moving stripes in the insect's visual field that would just bring about an orientation in the direction opposite to that of the movement. It was found that at low intensities the brighter of the stripes had to be nearly 100 times that of the lower in order to produce the response. This ratio decreased as the intensity was increased, and reached a minimum of 2.5, which was maintained at intensities 10,000 times higher than that at which the minimum was established.

Utilizing a similar experimental technique, Wolf (59, 60, 61) measured brightness discrimination in the honey bee. An observable response could be elicited from the bees between .01 and 100 millilamberts, and the thresholds of intensity discrimination were found to decrease with increasing brightness to a constant value which was maintained at the highest illuminations.

The curves of visual intensity discrimination for the clam, Drosophila, and the honey bee are presented in Figure 5.

2. *Vertebrates*

The most detailed work, in which brightness difference limens were obtained over a wide intensity range on infra-human forms, has been presented in the section above, and the present portion of the paper will consist mainly of an enumeration of the various experiments on vertebrates in which such thresholds were determined at only one or two brightness levels (in several of these experiments even the intensity of the standard stimulus was unknown or not given). This statement serves as an indication of the lack of information concerning visual intensity discrimination among subhuman vertebrates. A comparative summary of the differential

brightness sensitivity of vertebrates has been presented by the writer (Mead, 36).

Measurement of brightness discrimination among these forms has usually been accomplished by means of the Yerkes-Watson (64) apparatus or some modification of this discrimination technique (Lashley, 29; Smith, 50). Reeves (46) used such a modification of the discrimination technique in an aquarium containing several horned dace and sunfish. The dace learned to discriminate between the intensities of two experimental panels when the intensity difference was greater than a 1:4 ratio. Three sunfish discriminated a 1:2 ratio when tested according to the preference method, but could not be trained to make such a discrimination at this difference ratio.

Tugman (58) tested brightness discrimination in four English sparrows. The standard brightness was .098 candle power and the variable stimulus was .0025 candle power. Using the method of limits, the variable was gradually increased so as to approach the intensity of the standard and was then gradually decreased. The threshold (the difference between the least discriminable difference of the ascending and descending series) varied between .015 and .035 for the different subjects.

One of the earliest discrimination studies is that of Yerkes (63) on the dancing mouse. He found that a difference of one-tenth was sufficient to enable the animal to distinguish between two lights in the case of three standard values (5, 20, and 80 hefner units) of the stimulus. A detailed and precise study is reported by Moody (38) on brightness vision in the deer-mouse, this study being one of the few instances in which thresholds were found at a significant number of intensity levels. The average thresholds (expressed as the percentage of the discriminable differences to their respective standard intensities) for six subjects at 36.11, 40.84, 87.74, 156.5, and 493 candle power per $cm.^2$ were, respectively, 83, 87, 94, 95, and 109 per cent. Although there is a gradual increase in threshold at higher intensity levels, the author states that the Weber law seems to be upheld by his results.

Brightness difference thresholds have been determined in a number of instances for the rat, but generally in connection with other problems. Spencer (55), in his study of central inhibition, states that the white rat requires a difference of 70-80 per cent of a stand-

ard intensity in order to discriminate between two intensities. Lashley has measured the discrimination of light versus black in a number of studies with this animal, but only in one published study (28) have actual difference thresholds been found. In this experiment he notes that 13 of 17 normal rats discriminated lights whose difference ratio was 2:1. The intensity of the standard is not given, although it was probably about one millilambert. Using Lashley's (29) modification of the Yerkes-Watson box, Ghiselli (12) also found that the minimum ratio discriminated among 19 normal rats was 2:1, the average for all 19 animals being a ratio of 4.25:12. An approximately constant threshold of 32 per cent of a standard brightness was given by Slater (48) for this animal at four standard intensities. Munn and Slater (40) measured differential brightness vision in the albino rat at three illumination levels. When the standard was 3.20 millilamberts a threshold of 57 per cent of this standard was found for three animals. At 6.12 millilamberts one animal maintained the 80 per cent criterion of discrimination until the variable was 37 per cent of the standard, another animal failed to complete the series, whereas a third failed at a threshold of 46 per cent of the standard. The threshold for all three animals was 44 per cent of the standard when the latter was 19.29 millilamberts. Extremely consistent results are reported by Bills and Maukin (2) for the rat, but the brightness difference thresholds are expresseed in amperes so that comparison with other results is impossible.

In the raccoon, Munn (39) found the differential threshold for brightness to be somewhere between 33 and 21 per cent of a standard intensity at 13.45 millilamberts.

Smith (51, 52, 53) has been the only investigator previous to the present writer to secure intensity difference thresholds in the cat. The former used two standard intensities of .49 and 50.7 millilamberts, and also determined the effects of low and high surround illumination on this visual capacity. Using eight animals, this investigator reports threshold ratios of standard to variable of 1.23 to 1.6, the conditions of surround illumination not affecting the discrimination in the normal animals.

The Yerkes-Watson apparatus has been used by Stone (57) and Marquis (32) in studying brightness vision in the dog. The former used a standard intensity of one candle power (hefner units), and found that the smallest differences that two dogs could correctly

discriminate were, respectively, .14 and .20 candle power. These values correspond to ratios of 8.7:10 and 8.3:10. Human subjects, in the same apparatus, discriminated ratios of 9.0:10 and 9.17:10. Marquis (32) found essentially comparable thresholds for his normal animals; the dogs maintained the discrimination criterion until the stimulus panels differed in intensity by a ratio of 1:2, a value slightly inferior to that for human subjects in the same apparatus. An exceptionally keen threshold for brightness was obtained by Frolov (11), as reported by Razran and Warden (45), in the dog, using a conditioning method and reflected rather than transmitted light. This experimenter used a series of 50 papers whose intensities ranged from .1 to .5 candle meters. No. 50 was made the positive conditioned stimulus, and after a period of cruder differentiations No. 49 was finally discriminated from No. 50. The value of the Weber fraction for this threshold is .02, a ratio well below that reported for any animal below the human level.

Crawford (7) has found limens for brightness discrimination for the light adapted eye of four young Rhesus monkeys at four intensity levels. The animals were trained to reach toward the brighter window of a Yerkes-Watson light discrimination apparatus for food. The average Weber ratios for the four subjects at the four brightness levels, as judged by a criterion of 75 per cent correct responses, were as follows: .363 at .08 millilamberts, .098 at .77 millilamberts, .117 at 7.1 millilamberts, and .106 at 55.3 millilamberts. The average limens of two graduate students were slightly lower than the monkeys' at the upper three brightness levels and much lower at the lowest level. Although Crawford accepts this result as indicating a fundamental difference in the brightness sensitivity curves between the monkey and the human individual, the present writer suggests that such a decision should be held in abeyance until thresholds are determined over a wider range of illumination for both subjects in the same apparatus. Malmo[8] has found thresholds comparable to those obtained by Crawford with the monkey.

The curve of visual intensity discrimination of the human eye has been adequately determined. The earliest human curve was obtained by Aubert (1), but the classic data was obtained by Koenig and

[8]Personal communication.

Brodhun (26). Other human measurements have been made by Blanchard (3) and by Lowry (31), the latter devoting special attention to the form of the curve at high brightness levels. The curves obtained by these investigators all showed a tendency to rise at the higher illuminations, but the recent studies of Steinhardt (56) and Craik (6) indicate that when the eye is fully adapted to the testing illumination and when the surrounding illumination is the same as that of the test field, the Weber fraction remains constant and at a minimum at the highest intensities which can be used without painful effects. When these data are plotted on a double logarithmic plot there is a clear division into rod and cone function; "under these conditions the division almost leaps out of the data by itself" (Hecht, 22, p. 268). Rawdon-Smith (44), however, questions the applicability of a double log plot as an adequate description of all the human data. Steinhardt's (56) curve may be found in Figure 5.

The minimum thresholds found for the human eye are all superior to those obtained with any other organism. This superiority is undoubtedly partly due to the manner in which the thresholds are attained, i.e., with a bipartite test field, these conditions presenting an optimal state for brightness discrimination (Helmholtz, 24; Smith, 49). In those cases where the human threshold was measured under the same conditions as prevailed for the animals, although the human subject remained superior, the greater sensitivity was not as marked (Stone, 57; Marquis, 32; Crawford, 7).

C. Experimental

1. *Apparatus*

The present experimental arrangement is a modification of a lever discrimination apparatus suggested by Smith (50). His apparatus has been changed so as to permit observations on differential brightness discrimination in the cat.

In Figure 1 may be found the floor plan of the apparatus. Two discrimination boxes, 28 x 28 centimeters, are placed at one end of a table, 215 centimeters long. At the other end of the table is a light-proof restraining cage, 43 x 43 centimeters square. The distance from the restraining cage to the front of the discrimination boxes is 135 centimeters. A door in the front of each discrimination box may be

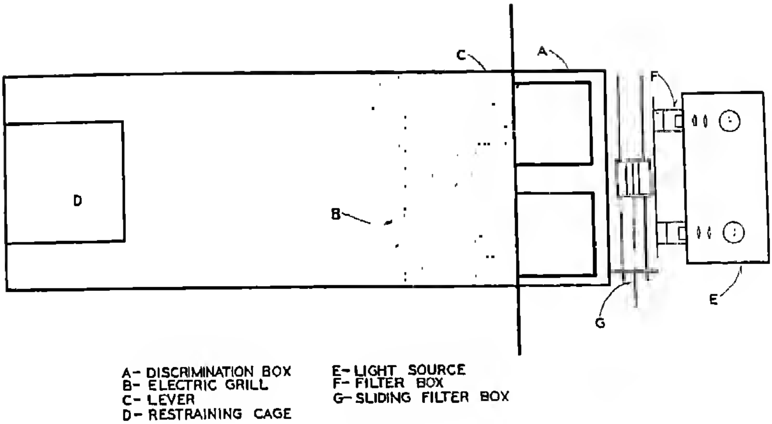

FIGURE 1

FLOOR PLAN OF BRIGHTNESS DISCRIMINATION APPARATUS FOR THE CAT

opened by the depression of a lever which extends from the lower front of each box. In the doors of the two boxes are inserted panels of flashed-opal glass. Also in the door, and in front of the glass panel, is placed a black cardboard diaphragm containing a 10 x 10 centimeter hole in its center, thereby leaving an expanse of glass of the same dimensions as the opening visible from in front of the discrimination boxes. The distance between the centers of the visible glass panels is 38 centimeters. The rear of each discrimination box is cut away so as to permit the passage of a light beam and also the insertion of a frosted glass panel.

The animal is released from the restraining cage by raising a sliding door on the front of the cage by means of a cord which extends behind the experimenter's screen (not shown in Figure 1). This screen stands four feet away from and along the side of the large table containing the discrimination boxes. A small window cut in this "Kluge Hans" screen allows observation of the animal's activity during experimentation. After release from the restraining cage (*D*) (in Figure 1) the animal discriminates between the 10 x 10 centimeter light panels by depressing the lever (*C*) in front of the brighter panel. This causes the door of one of the discrimination boxes (*A*) to open, and the cat obtains a small amount of fish placed inside the box. Food was always kept in both boxes during experi-

mental sessions, thereby avoiding olfactory cues. The animal may be punished for an incorrect choice by means of an electric shock delivered through the levers (*C*) and the wire grid (*B*) placed in front of the discrimination boxes. The intensity of the shock is governed by a small transformer which was adjusted for each animal so as to be just "disagreeable." Shock intensity was generally of the order of 1-2 milliamperes. A small neon lamp connected in series with the shock circuit, and located behind the experimenter's screen, indicates that the subject has made an incorrect choice.

As a check against the possibility that the subjects might obtain a shock cue which was not "detected" by the neon lamp, a device was installed which required the animal to actually depress, rather than just touch, the lever before electrical contact was made.

The light sources are located in another box (*E*), 30 x 60 centimeters, placed on a table behind the discrimination boxes. The filament of each lamp lies in the same horizontal and vertical plane as the centers of the discrimination panels. Lenses placed in the beam of each light serve to narrow the width of the beam which is projected through circular holes, 6 centimeters in diameter, in the front of the source box. Appropriately placed water cells aid in reducing the quantity of heat which is radiated from the lamps. Each light beam then passes through and comes to a focus in a small filter box (*F*), and its intensity may be cut down to any desired amount by means of Wratten neutral tint filters inserted at this point. The light rays then pass through the back of the respective discrimination boxes and are diffused by the flashed-opal glass panels located in the doors of these boxes. A large screen, extending over the edge of the large table in the plane of the doors of the discrimination boxes, effectively eliminates all stray light from the rear of the apparatus so that the brightness of the stimulus panels are the only sources of illumination when the boxes are viewed from the front. All parts of the apparatus in the visual field of the subject, except the stimulus panels, are painted flat black.

Either light beam may be interrupted by any desired combination of filters placed in the sliding filter box (*G*), movement of which permits the position of the brighter (positive) stimulus panel to be reversed from left or right. The filters used in this box were 4 x 5 inches in size. Some were differentially exposed Eastman process

plates, and others were pieces of plain glass. The transmission power of each of these large filters, as well as that of various combinations of them, was calibrated on the basis of readings taken from the stimulus panels with a MacBeth illuminometer.

The brightness of the stimulus panels is controlled by using different wattage lamps as light sources when testing is done at brightness levels of (ca.) 1 millilambert and higher. Lower brightness levels are produced by using 6-volt lamps as sources and reducing the resulting brightness with the Wratten filters.

It should be noted that vibratory tactual and kinesthetic cues which might have accompanied movement of the sliding filter box were avoided, this box being on a table separate from that on which the discrimination boxes are located.

2. *Animals*

Three normal male cats (Subjects 1, 2, and 3) were trained in the brightness discrimination habit. The ages and past history of these subjects are unknown. With the exceptions mentioned below, the health of all the subjects remained excellent during 3½ years of experimentation. Their diet consisted of balanced dog food rations to which was added a small amount of cod liver oil daily. They were ordinarily "run" just before receiving their daily ration of solid food so that they were "24 hours hungry." They received, however, a small quantity of milk at a time usually mid-way between successive experimental periods.

3. *Procedure*

Each animal was first made familiar with the apparatus by being fed in the discrimination boxes. The doors to these boxes were then closed and training on the lever-pressing habit was begun. The stimulus panels were not illuminated during this period of preliminary training. After the lever habit was established the stimulus panels were equated at an illumination of 300 millilamberts and a filter of 1/40th transmission was placed in the movable filter box. The animals were then trained to select the brighter of the two stimulus panels. The position of the positive (brighter) stimulus panel was shifted from right to left in chance order during training on the brightness discrimination habit thereby avoiding establishing a preference habit during training. When making threshold deter-

minations the positive stimulus was changed from left to right in a random way except with the restriction that it never occur more than three times in succession on the same side. An exception was made to this rule when the animal demonstrated a "position habit." In this event the brighter stimulus was presented successively on the side opposite to that of the position habit until the habit was broken. Position habits seldom appeared, however, until the brightness difference approached a liminal value. The subjects were considered as having learned the intensity discrimination habit when they first correctly chose the brighter panel for 25 successive trials (one experimental period).

In making threshold determinations the beginning trials were always made with a relatively large brightness difference between the stimulus panels, i.e., the filters in the movable filter box were of low transmission. A criterion of 76 per cent correct responses in 25 trials was selected as indicative of a clearcut brightness discrimination. In other words, if the subject made seven or more errors in the daily period of 25 trials he was considered as failing to discriminate that particular brightness difference. If the animal did 76 per cent or better on a particular day, the brightness difference was decreased on the following day by increasing the transmission of the filters in the movable filter box by a small amount. Ordinarily the transmission of these filters at the beginning of a threshold series was about 11 per cent (corresponding to a Weber fraction of 8.09) and this was gradually increased, following each day's successful performance, in steps which averaged about 7 per cent. The highest transmission ever used was 95 per cent, corresponding to a Weber fraction of .05. If an animal failed to meet the discrimination criterion for six successive experimental periods the threshold for that particular brightness level was taken as the brightness difference which the cat last "passed." For the first 18 months of experimentation each subject was run for 25 trials a day regardless of the number of errors made. After this time the animals were tested each day until they made seven mistakes or had made 25 comparisons. The method is, then, a modified method of constant stimuli.

When obtaining thresholds at standard intensities of less than one millilambert the cats were dark-adapted for at least 30 minutes before beginning the experimental period.

About 400-500 trials were required to obtain a differential bright-

ness threshold, and after the first few limens were determined it was found that each subject "transferred" readily to other intensity levels. Thus, although the subject during the completion of the last threshold had been performing with less than 76 per cent accuracy, when a new standard intensity and a large brightness difference were introduced, the cats immediately showed a high percentage of accuracy, generally above 88 per cent.

Control runs were occasionally instituted. They consisted in carrying out the procedures listed above but without placing any filters in the movable filter container. There was, therefore, no brightness difference for the animals to discriminate. In these instances the scores all approximated that which would be expected by chance. The highest score obtained under these conditions for any one animal was 64 per cent correct responses in 25 trials.

4. *Results*

The experimental findings will be presented under two main headings: (*a*) the number of trials required to set up the brightness discrimination habit, and (*b*) the thresholds of brightness discrimination as a function of test illumination. Some of the present results have already been reported (Mead, 34, 36, 37).

a. Learning the brightness discrimination habit. It will be remembered that that criterion of learning used was one errorless experimental day, or 25 consecutive errorless trials. Furthermore, the performance acquired by the cats in the present study consists of at least two stages, namely, the lever-pressing habit and the brightness discrimination habit. There was no brightness difference between the stimulus panels during the training on the lever, and the scores that will be reported below are the number of trials required by the animals to learn to select the brighter panel after they had learned to press the levers to obtain entrance to the food boxes.

The learning data for the three cats discussed above, and for two other normal male cats trained under similar conditions in the same apparatus, are presented in Table 1. Subject 1, the best motivated animal of the group, learned to distinguish between the intensities of the stimulus panels in approximately 100 trials, or four experimental periods. Subjects 2 and 3 required 350 trials before they were able correctly to choose the brighter panel for 25 consecutive trials. Shortly after the training of these last two subjects was begun, they suffered a severe cold and the training was temporarily inter-

TABLE 1
TRIALS REQUIRED BY FIVE NORMAL CATS TO LEARN THE BRIGHTNESS DISCRIMINATION HABIT

Cat No.	Brightness level	Brightness difference	Trials
1	300 ml.	40:1	100
2	300 ml.	40:1	350
3	300 ml.	40:1	350
4	2.08 ml.	40:1	185
5	.128 ml.	10:1	200

rupted. This fact may account for their slower rate of learning. Although subjects No. 4 and No. 5 were trained at lower brightness levels, and, in the case of the latter animal, at a different brightness difference, they reached the criterion of learning for the visual discrimination habit in 185 and 200 trials, respectively. They thus appear to fall at a point approximately mid-way between the total number of trials required by Subject 1 and Subjects 2 or 3. The average for all five animals is 237 trials, and this may be spuriously high due to the inclusion of the data for Subjects 2 and 3.

Despite the fact that the learning criterion was more strict than that customarily found in discrimination studies, the animals employed in this experiment achieved the learning criterion in a number of trials which is representative for this type of performance. It would thus appear that the apparatus utilized here permits preliminary training in a reasonably efficient period of time.

b. The curve of visual brightness discrimination. Thresholds of intensity discrimination have been found for Cats 1, 2, and 3 over a wide range of intensity levels. These limens are presented for the respective animals in Tables 2, 3, and 4, and correspond to the difference limens obtained at each level of illumination used in testing.[4] The tables represent, respectively, the threshold capacities of Subjects 1, 2, and 3. The column at the extreme left presents the brightness level of the standard stimulus. The threshold (Weber ratio) corresponding to this intensity is found in the fourth column from the left. The other columns are the mean intensities at log-unit steps (second column from the left), the corresponding mean Weber ratios

[4]The writer considers it unnecessary to present tables showing the performance of each cat at each decrement in brightness difference for a given threshold series. Most of these data may be found, however, in the writer's (35) doctorate thesis which is on file at the University of Rochester Library.

TABLE 2

THRESHOLDS FOR SUBJECT 1

Brightness (ml.)	M_B	Log M_B	$\frac{\Delta I}{I}$	$M_{\frac{\Delta I}{I}}$	Log $M_{\frac{\Delta I}{I}}$
.000000865 .000000132	.000000499	—7.70	.82 ∞	∞	∞
.00000132	.00000132	—6.12	1.04	1.04	.02
.0000108 .0000132 .0000234	.0000158	—5.20	.52 .49 .59	.53	—1.72
.000135 .000128	.000132	—4.12	.52 .41	.47	—1.67
.00293 .00256	.00275	—3.44	.24 .24	.24	—1.38
.0211 .0984	.0598	—2.78	.45 .19	.32	—1.51
.210 .126 .129 .153	.155	—1.19	.35 .19 .19 .19	.23	—1.36
2.06 3.96 3.53	3.18	.50	.21 .38 .38	.32	—1.51
37.5 26.1	31.8	1.50	.21 .19	.20	—1.30
214.5 150.0 158.0 176.0 323.0	204.3	2.31	.12 .24 .24 .24 .19	.21	—1.32
1397.0 3977.0	2687.0	3.43	.41 .32	.37	—1.57

(fifth column from the left), and the logarithmic values, respectively, of these means (third and sixth columns, respectively, from the left). In these tables the thresholds are presented in order of increasing brightness level of the standard stimulus. This order is not the same as the sequence of brightness levels which was followed empirically. In fact, the experimental arrangement of successive brightness levels was such that a jump of two or more log unit steps was generally made.

In examining the trends in the data and in comparing one subject

TABLE 3

THRESHOLDS FOR SUBJECT 2

Brightness (ml.)	M_B	Log M_B	$\frac{\Delta I}{I}$	$\frac{M_{\Delta I}}{I}$	Log $\frac{M_{\Delta I}}{I}$
.000000865	.000000865	—7.94	3.34	3.34	.52
.0000108	.0000108	—5.03	1.22	1.22	.09
.000135	.000135	—4.13	.72	.72	—1.86
.0126 .0139	.0133	—2.12	.38 .52	.45	—1.65
.185 .120 .153 .153	.153	—1.19	.38 .52 .52 .52	.49	—1.69
2.06 3.96 3.53	3.18	.50	.61 .82 .52	.65	—1.81
37.5 26.1	31.8	1.50	.45 .72	.58	—1.76
214.5	214.5	2.33	.45	.45	—1.65

with another it is probable that the comparisons will be more clear if we consider the mean $\frac{\Delta I}{I}$ column in Tables 2, 3, and 4. It will be observed that for each subject the differential brightness threshold decreases from a relatively high value at low intensities to a much lower value at high illuminations. When the positive stimulus is at an extremely low value, and when this cannot be differentiated from a black negative stimulus, it becomes possible to secure a measure of the absolute brightness threshold. This was found, for Cats 1 and 3, to be in the vicinity of $.132 \times 10^{-6}$ millilamberts. Discrimination at this level of the positive stimulus was extremely poor (Weber fraction of infinity) for these animals. Discrimination was only slightly improved at the next higher brightness level ($.865 \times^{-6}$ millilamberts) for Cats 1 and 2. A rapid improvement was noted when the level was raised still further by one and two log unit steps. Thus subjects 1 and 3 performed much better at $.132 \times^{-5}$ millilamberts and Subject 2 improved at 1.08×10^{-5} millilamberts. The mean thresholds obtained at the above-mentioned brightness levels are, however, all inferior to those obtained at any higher level of illumi-

TABLE 4

THRESHOLDS FOR SUBJECT 3

Brightness (ml.)	M_B	Log M_B	$\frac{\Delta I}{I}$	$\frac{M_{\Delta I}}{I}$	Log $\frac{M_{\Delta I}}{I}$
.000000132	.000000132	—7.12	∞	∞	∞
.00000132	.00000132	—6.12	2.84	2.84	.45
.0000108 .0000132 .0000234	.0000158	—5.20	2.03 .59 .85	1.16	.07
.000135 .000128	.000132	—4.12	1.03 .59	.81	—1.91
.00865 .00293	.00579	—3.76	.72 .49	.62	—1.79
.0211 .0984	.0598	—2.78	.75 .47	.61	—1.79
.210 .126 .139 .153	.157	—1.20	.61 .52 .38 .52	.51	—1.71
3.50 3.96 3.53	3.67	.57	.61 .72 .72	.68	—1.83
26.1	26.1	1.42	.52	.52	—1.72
214.0 150.0 158.0 176.0 323.0	204.2	2.31	.45 .59 .41 .41 .47	.47	—1.67
1397.0 3977.0 3583.0 3321.0	3070.0	3.49	.59 .47 .61 .47	.54	—1.73

nation. The greatest drop in the brightness discrimination curve therefore appears to occur between the absolute threshold and 1.08×10^{-5} millilamberts, a range of two log unit steps. Cats 1 and 3 continued to improve at this level over their performance at lower levels, but their mean threshold at this illumination is still inferior in differential sensitivity to that found at all higher illuminations.

Performance continues to improve as the intensity of the positive stimulus is increased, but the rate is irregular. This irregularity makes it difficult to state exactly at what illumination differential sensitivity is best for the cat. For Subject 1 a low Weber fraction

is first established at $.275 \times 10^{-2}$ millilamberts, the lowest mean threshold occurs at 31.8 millilamberts, whereas the keenest single threshold was found at 204 millilamberts. Subject 2 discriminated best at 1.33×10^{-2} and at 215 millilamberts. The first minimum for Subject 3 was found at 15.7×10^{-2} millilamberts whereas the lowest Weber ratio for this animal was established at 204 millilamberts. At the extreme high intensities (log I equals 3.00 and above) the performance of Cats 1 and 3 was slightly inferior to their discrimination ability at lower intensity levels. (It was unfortunately impossible to observe the performance of Cat 2 at this high brightness value because the animal died of distemper during a summer vacation. The data obtained with this animal were gathered over a period of 20 months).

Figures 2, 3, and 4 graphically depict the data presented above.

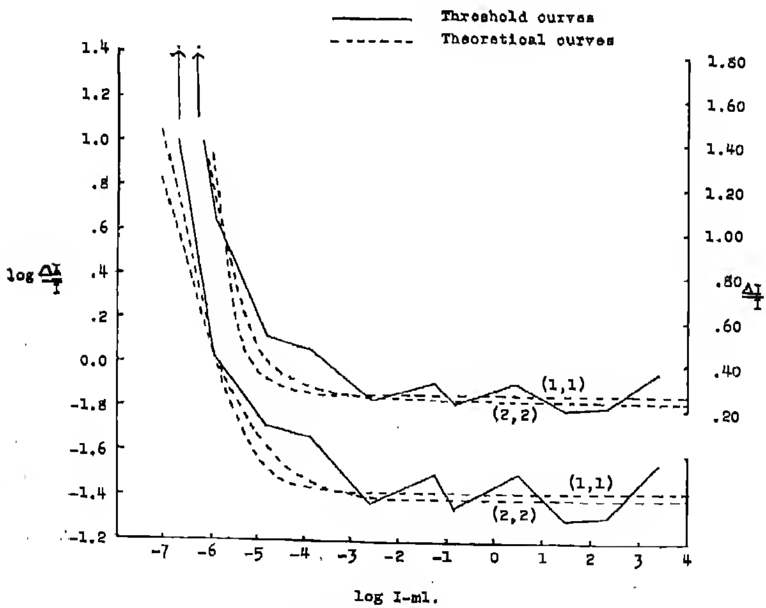

FIGURE 2

Mean Thresholds at Averaged Log Unit Steps and Fitted Theoretical Curves of Cat 1

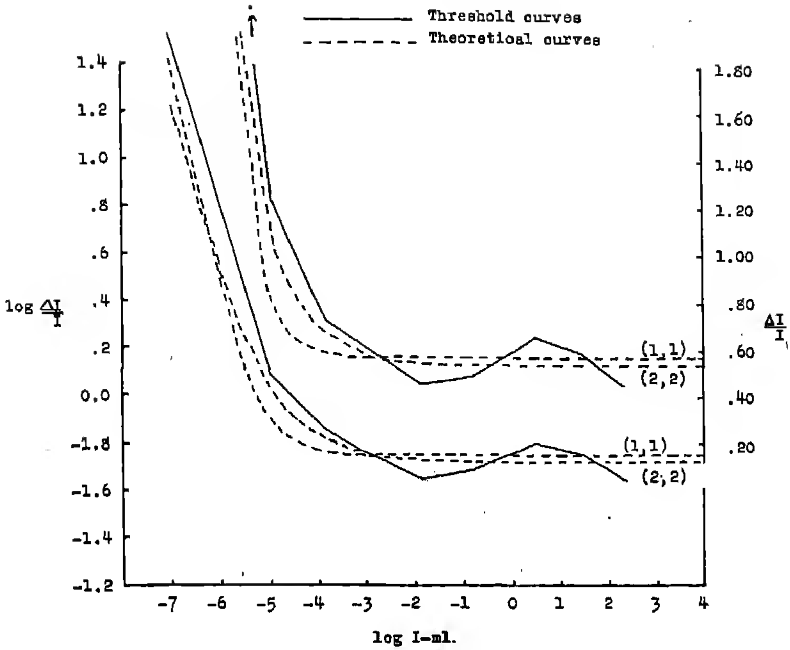

FIGURE 3
MEAN THRESHOLDS AT AVERAGED LOG UNIT STEPS AND FITTED THEORETICAL CURVES OF CAT 2

In these figures the solid lines are drawn through the mean Weber ratios at averaged log unit steps (second column from the left in Tables 2, 3, and 4). The lower solid line represents the logarithm of the mean Weber ratios (sixth column from the left in Tables 2, 3, and 4) whereas the upper solid line is drawn through the mean Weber ratios themselves (fifth column from the left in above tables). The double- and semi-logarithmic plots are used because of Hecht's (22) suggestion that the former emphasizes the rod-cone shift in the human data. Values for the logarithm of the Weber fraction are given on the ordinate at the left and the absolute values for this fraction are shown on the ordinate at the right.

These figures clearly demonstrate the changes in the mean Weber ratio as a function of illumination. The fact that this ratio is infinitely great at the absolute threshold is indicated by the break in

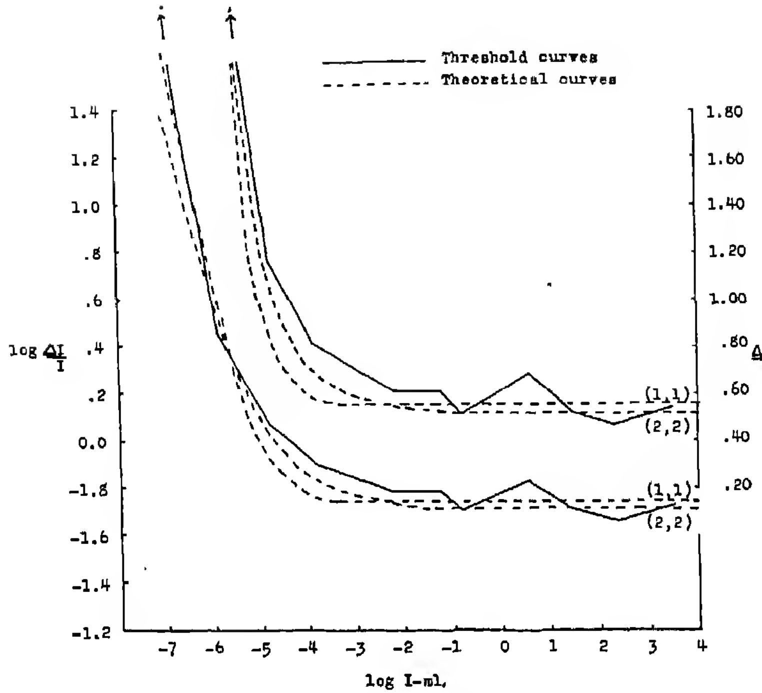

FIGURE 4

Mean Thresholds at Averaged Log Unit Steps and Fitted Theoretical Curves of Cat 3

the solid curves. Improvement in differential sensitivity is rapid as the illumination level is raised, and the betterment in performance is maintained for all three subjects up to a standard intensity level of approximately 1.00×10^{-4} millilamberts. With increasing intensity, although the improvement continues, there is considerable irregularity in the curves. The reader will note that there is some increase in the height of the brightness discrimination curve at the greatest illumination levels for Subjects 1 and 3. Individual differences among the animals remain fairly consistent. Thus Subject 1 did better at all brightness levels than the other two subjects, who seem to be fairly equal in this visual capacity.

A question which arises as the above trends in the data are examined is the significance of the changes in the limens as the illumina-

tion level is varied. This question is especially cogent at intensities higher than that at which the first minimum in the thresholds is established. From our knowledge of the curve of intensity discrimination in the human individual (Hecht, 22; Steinhardt, 56), we might expect in the cat to find a break in the curve, thus indicating a shift from rod to cone function as we ascend the intensity scale. Rawdon-Smith (44), however, has suggested that the shift from rod to cone vision may be gradual over a brief range of intensities when vision is "mixed." In the present case these speculations are difficult to answer since the variability in thresholds which have been found at approximately equal illuminations is such as to obscure any real break in the intensity discrimination curve.

Since one of the purposes of the present investigation was to test the generality of Hecht's (19, 20, 22) theory of brightness discrimination we may now ask how well does the above data fit his theoretical description of this curve. It will be remembered that his family of theoretical curves are four in number, each with the same limiting values at the highest and lowest illuminations, but with varying slopes at intermediate brightnesses. The limiting curves are for those cases where the order of molecularity of the photochemical and dark reactions are both equal, respectively, to 1 and to 2. It was thought at the outset of this experiment that the best-fitting theoretical curve for the findings with the cat could be determined by using a double logarithmic plot of the data and then selecting the most probable theoretical curve by simple inspection. The variability in the experimentally determined curves makes this an unwise procedure, however, so the obtained thresholds were fitted to the theoretical equations by the method of least squares. This was done for the limiting cases mentioned above; i.e., when m and n are equal to 1 and where m and n are equal to 2, m representing the order of molecularity of the light reaction and n that of the dark reaction.

Another question which arises during the curve-fitting process is whether there is a rod-cone shift in the data, and, if so, should the data be broken into two portions and separately fitted to the theoretical curves. The suggestion of such a shift is made when we examine the empirical curves for Cats 2 and 3 in Figures 3 and 4. In these figures the solid curves may be seen to fall to a minimum as the brightness level is raised from the absolute threshold, there is then a rise in the curve at median illuminations followed by another fall before the highest intensity level is reached. This contour of

the brightness discrimination curve is not clear-cut for the curve of Cat 1, however, as indicated by the solid curves in Figure 2. It was therefore concluded that the evidence for a shift in discrimination sensitivity in the cat at scotopic and photopic levels is equivocal in the present experiment, and the data were not broken into two separate portions for the curve-fitting procedures.

The broken lines in Figures 2, 3, and 4 respectively are the best-fitting theoretical curves for the limiting cases of m and n of Cats 1, 2, and 3, respectively. The numbers in the parentheses stand for the order of molecularity of the photochemical and dark reactions. Both double- and semi-logarithmic plots of these best-fitting theoretical curves are shown, as was the case for the raw data. The fact that a given pair of theoretical curves do not meet at the extremes of intensity is partially due to the rounding-off of figures in the mathematical computations. The theoretical curves for Cat 1 in Figure 2 resulted when 27 different thresholds and their corresponding brightness levels were substituted in Hecht's limiting equations for the brightness discrimination function. The same procedure was carried out with 15 thresholds for Cat 2 (see Figure 3) and with 27 thresholds for Cat 3 (see Figure 4).

On the basis of the above results it is apparent that no single one of Hecht's family of theoretical curves can be used to describe the brightness discrimination curve of the cat. This may be partially due to the magnitude of the step-changes in the brightness differences used in obtaining thresholds. By casual inspection the curve where m and n equal 2 seems to best describe the empirical data for all subjects between brightness of 1.00×10^{-6} and 1.00×10^{-3} millilamberts, the curve where m and n equal 1 best apply at brightnesses less than 1.00×10^{-6} millilamberts, and either of these curves might apply at brightnesses greater than 1.00×10^{-3} millilamberts. The general contour of the data does, however, follow the theoretical expectations, and does resemble the curves obtained for other invertebrate organisms and for man. The empirical results for the cat and these other forms demonstrate that brightness discrimination is poor at low levels of intensity but increases rapidly as the level is raised. It is probable that the curve for the cat would also remain at a constant level at the highest illuminations if full adaptation to the test lights had been permitted and the surround illumination equated to the test patches.

D. Discussion

1. *Learning and Retention of the Brightness Discrimination Habit*

In the foregoing pages a method has been described which permits the determination of thresholds of visual intensity discrimination in the cat as a function of changing illumination. Some evidence as to the comparability of the described method to that of other investigations of animal vision may be gained by a comparison of the learning scores of the present subjects with those obtained on comparable organisms using different experimental arrangements. The present apparatus allowed quick and easy changes in the position of the positive stimulus and also preserved the desirable features of prior experimental methods.

The number of trials required by five normal cats to learn the brightness discrimination habit to a criterion of 100 per cent correct responses in 25 successive trials compares favorably with the number of trials required by the cat in other similar experimental situations (Smith, 53; Smith and Kappauf, 54). The average for all five animals was 237 trials. This rate of learning the brightness discrimination habit reported here for cats is similar to that reported in the rat (Lashley, 28) and the dog (Marquis, 32, 33), although in the studies on the rat the training was carried out in a light-dark situation. It may be that experiments which have utilized this condition of training offer a somewhat more easy learning situation for the animal. Marquis (32) found that the average number of trials required for 6 dogs to reach a learning criterion of 100 per cent for 30 successive trials was 160. The average for the five normal cats reported above would undoubtedly have approached this figure had not the training of two of them been interrupted by illness.

After the brightness discrimination habit has been well established in the present apparatus, the cat retains the habit over a long period of time. Two of the animals reported on above were not run for as long a period as six months and yet were able to meet the discrimination criterion during the first experimental session after this rest period. One of these cats manifested 92 per cent correct responses in this 25 trial session. Reference should also be made to the ease with which the animals transferred from one brightness level to another, indicating that the subjects were responding to relative rather than absolute differences in stimulus brightness.

2. *The Threshold of Brightness Discrimination*

a. Variability in thresholds. The limens obtained in this study at different brightness levels deviate to some degree from the best-fitting theoretical curves. Furthermore, in those cases where difference limens were obtained twice or more at approximately the same intensity level, there is some dispersion around the mean Weber ratio. This lack of reliability may be accounted for in a number of ways, among which we may mention: (*a*) The length of time required to establish the threshold. Each liminal value involved three weeks or more of experimentation. It is not unlikely that during this time the physiological state and motivating conditions of the subjects varied, and thereby contributed to variations in the keenness of discrimination. (*b*) The discontinuity of the changes in the brightness difference involved in the incremental steps of filter transmission. The average change in filter transmission used in going from one brightness difference to another was 7 per cent, but this corresponds to a relatively large difference in the Weber fraction computed for these two brightness differences. Thus if the subject passes the discrimination criterion at two different times when the brightness difference varies by only one experimental step, there is nevertheless an apparently large change in the threshold ratio. (*c*) The distance between the test fields. With human subjects an increase in the separation of the test fields produces a marked decrease in differential sensitivity, and markedly increases the difficulty in making such a judgment. The wide separation between the test fields in the present apparatus may thus be contributing to the apparent inconsistency of the results.

The reliability of the limens measured at any single brightness level appears to be increased if we consider the observed variability in terms of the number of step-changes of the variable (negative) stimulus used in the threshold series. An examination of the thresholds (Figures 2, 3, and 4) obtained after the intensity discrimination curve has flattened out reveals that, upon retesting, the liminal variation at any one log unit of brightness is of the order of one or two incremental transmission changes of the filters used in changing the variable stimulus. Considering the crudeness of measurement of these differential intensity limens in the cat (as compared to that used in testing the human individual), the curves obtained in the present experiment appear to be a valid measurement of this visual capacity, and the definite trend observed in the Weber fraction at different

illuminations appears to be a reliable indication of the manner in which differential sensitivity changes as a function of illumination in this animal.

A final point which may be considered in this section is that variability is probably existent to some extent in all determinations of visual capacity with infra-human animals. The writer knows of no studies of brightness discrimination in such organisms in which the reliability of obtained limens was checked by retesting. The thresholds reported in the literature for a number of different mammals are based upon a single liminal series after only a short period of training in the discrimination. We may use the finding of variability upon retesting in the present investigation as a basis for the suggestion that future studies on infra-human brightness discrimination should take some measure of the reliability of obtained limens. The question of reliability is especially significant in evaluating pre- and post-operative thresholds and their divergence in relation to the effect of surgical modification of certain neural centers of the visual system. It is believed that the data presented above for the cat may be utilized as an adequate index of the capacity of the normal cat in brightness discrimination, and may furthermore serve as a basis for judging defects in this type of vision produced by lesions in the central nervous system.

b. The absolute threshold. By the present method it has proved possible to measure the least perceptible brightness to which the cat is capable of responding. When the standard (brighter) stimulus is decreased to such a level that the cat fails to distinguish it from zero illumination on the negative stimulus panel it may be assumed that the former has fallen below the absolute brightness threshold. This threshold was found to be in the vicinity of 1.3×10^{-7} millilamberts for Subjects 1 and 3, which is within one log unit of the average value reported by Bridgman (4) for six normal cats.

In answer to the belief that cats see better in the dark than humans we may posit a tentative answer from our knowledge of this threshold. From his review of the literature Bridgman (4) has stated that the most representative figure for this limen in the human individual is 5.8×10^{-7} millilamberts. A comparison of his threshold for the cat and human indicates that the human is less sensitive by an approximate factor of ten. However, as pointed out by Bridgman, the comparatively high sensitivity of the cat's eye can be explained by the fact that the "speed" of the latter's optical system and reflection

by the tapetum (Murr, 41) permit a greater amount of light from a given source to fall on the cat's retina than on the retina of the human subject. Thus it is unlikely that the retinal sensitivity of the human subject is any less than that of the cat albeit the latter organism can respond to a visual stimulus of lower brightness value. While making the absolute threshold measurements for the subjects in this study, the writer was completely unable to see either the cats or the stimulus panels even after a considerable period of dark adaptation.

c. The difference threshold. A phylogenetic comparison of differential brightness thresholds has already been reported (Mead, 36). Subject 1 showed the keenest sensitivity throughout the entire experiment, and the limens for this animal were taken as representative for the cat. The smallest Weber ratio obtained with this subject was .12 at a brightness of 215 millilamberts. Since this ratio was found only once, whereas the next higher ratio of .19 was obtained six times, the more reliable conclusion was reached that the maximum sensitivity of the cat corresponds to a Weber ratio of .19. Weber fractions of this low order were obtained between 9.84×10^{-2} and 323 millilamberts.

Comparing those values obtained in studies utilizing a training method and in which the stimulus panels were simultaneously presented in the visual field of the subject, the Weber thresholds[5] for brightness discrimination of various infra-primate mammals are found to correspond very closely with those obtained here for the cat. The minimum threshold reported on a reliable number of subjects for the mouse is .83 (Moody, 38); for the rat, 1.0 (Lashley, 28; Ghiselli, 12), .59 (Munn and Slater, 40), and .43 (Spencer, 55); for the raccoon, .27 (Munn, 39); and for the dog, .16 (Stone, 57) and .20 (Marquis, 32). Smith (53) reports a threshold difference corresponding to .23 at two brightness levels for the cat. A comparable study (Crawford, 7) on the monkey has yielded a threshold of .098. The cat is thus seen to equal or excel many of these other animals. The thresholds reported in the present study are superior to those found for the rat, in the general range of those previously reported for the dog and raccoon, and inferior to those obtained for the monkey. There is, therefore, a type of relation existing between differential brightness sensitivity and phylogenetic level. The physiological foundation of these differences is probably traceable to numer-

[5]Computed from the original experiments by the present writer.

ous differences in the peripheral visual system, which must yet be ascertained for specific comparisons.

d. The curve of brightness discrimination. The primary purpose of the present investigation was the determination of the curve of brightness discrimination in the cat as a function of illumination. Similar data for the human eye (Steinhardt, 56), the clam (Hecht, 17), Drosophila (Hecht and Wald, 23), and the honey bee (Wolf, 59) have been shown by Hecht (19, 22) to be adequately described by the latter's theoretical curves based on his general photochemical theory of intensity discrimination. These best fitting curves are shown in Figure 5. The numbers in the parentheses represent the order of molecularity of the light and dark reactions and refer to the theoretical equation which best describes the data for each organism. The best fitting curve for the above organisms was determined by using a double logarithmic plot of the empirical data and comparing it visually with a similar plot of the theoretical equations. Variability in the cat data makes this method of curve fitting an unwise procedure for the present results. These data were therefore fitted to the limiting cases ($m=n=1$ and $m=n=2$) of the theoretical equations by the more complicated method of least squares. The resulting curves are shown as dashed lines in Figure 5 and are for the "average" cat, i.e., for the average of the Weber ratios at averaged log unit steps of Subjects 1, 2, and 3. The best fitting theoretical curves for the individual animals have already been presented in Figures 2, 3, and 4. The experimentally observed data for the average cat is represented by a solid unsmoothed curve in Figure 5.

If we take the points on the theoretical curves in Figure 5 where each organism is most sensitive (lowest Weber ratio) to brightness differences, and use absolute rather than logarithmic value of $\frac{\Delta I}{I}$, and employ the human curve as the standard, it is found that Mya is 570 times less sensitive than man, the fruitfly inferior by approximately 150 times, the bee inferior by about 40 times, and the cat inferior by a factor of about 20. The absolute brightness threshold for the cat, however, is lower than that of all these other forms.

The empirical observations gathered to date for the cat are not sufficient to establish critically the proper orders of molecularity for the light and dark processes in the theoretical equations to which the data was applied. That is, no *single* one of Hecht's family of

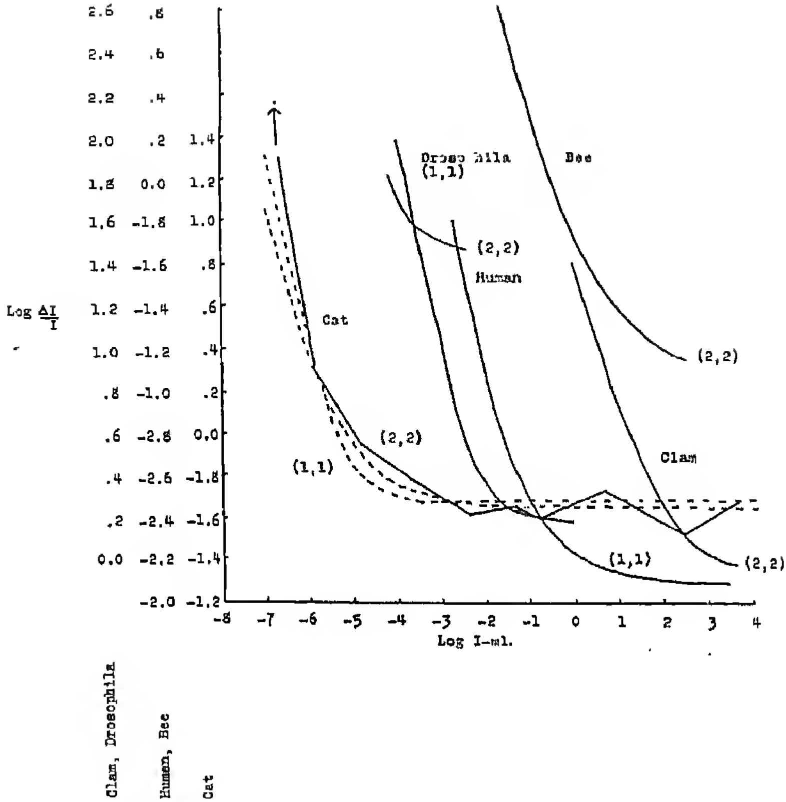

FIGURE 5
BEST-FITTING THEORETICAL CURVES OF INTENSITY DISCRIMINATION FOR VARIOUS ORGANISMS COMPARED TO THE AVERAGE CAT

theoretical curves can be said to best fit the results on the cat. However, the data as a whole are approximately described by Hecht's theoretical curves, and since these are based on a consideration of a general theoretical photochemical description of brightness discrimination we may emphasize the degree of similarity of the obtained data to the theoretical expectations. The Weber ratio, as predicted theoretically and shown experimentally, for the cat is an inverse function of the illumination level at which the threshold is determined.

The human curve of brightness discrimination shows a break which coincides with the taking over of function by the retinal cones (Hecht, 22), the rods contributing to the low luminosity portion of the curve. Since the cat's retina probably contains both rods and cones (Chievitz, 5; Krause, 27; Murr, 41, 42; Schultze, 47; Zürn, 65), it may be asked why there is not a similar break in the intensity discrimination curve for this animal. The present evidence is not clear-cut on this point. There is some suggestion of a shift for Subjects 2 and 3, but this may be due to a number of experimental variables which may serve either to augment or obscure a sudden change in the curve. The factors producing variation in obtained thresholds at similar luminosity values have already been discussed. A rise in the curve may be produced if the eye is not completely adapted to the test field illumination (Steinhardt, 56; Hecht, 21; Craik, 6). Brightness discrimination in the human eye is also impaired if the surround illumination is not the same as the test object (Steinhardt, 56; Craik, 6), although Smith (51, 52, 53) found that a high and low intensity of surround illumination produced no obvious change in discrimination acuity in the normal cat. Rawdon-Smith (44) has emphasized that the change from rod to cone vision as the brightness level is shifted from scotopic to photopic illuminations may be a gradual one. Indication of a rod-cone shift in visual discrimination by the cat has been found between 10 and 100 millilamberts by Kappauf (25) in his flicker study with this animal. It would appear, therefore, that the crudeness of the step-changes in brightness difference which were used in all threshold series, plus the factors just mentioned, may have obscured the appearance of a rod-cone break in the brightness discrimination curve for the cat.

To summarize the contribution of the present paper to our knowledge of intensity discrimination, the present observations show that a theory which has been found to apply to the data for this visual function in man and invertebrate animals are here shown to approximately apply to the same visual capacity in the cat. Since Hecht's theoretical curves appear to describe the data on intensity discrimination for the cat, it may be said that his theoretical formulations are given wider significance. The present data offer support for the assumption that his theory applies to a large portion, if not all, of the phyletic series.

E. Summary and Conclusions

An apparatus and method have been described which permit the measurement of differential brightness thresholds in the cat. Five normal cats have been trained to make such responses in this apparatus, and a large number of visual intensity discrimination limens over a wide range of illuminations have been secured on three animals. The results and conclusions of the present study may be summarized as follows:

1. The average number of trials required by five cats to reach a learning criterion of 25 consecutive errorless trials was 237. One animal learned the brightness discrimination habit in 100 trials whereas two animals required as many as 350 trials to meet the criterion of learning.

2. Incidental observations indicate that cats which have been well trained in the present apparatus retain the discrimination habit over at least a six months period of nonexperimentation.

3. The sensitivity of the cat in this type of visual discrimination is found to equal or excel that of other infraprimate mammals. The lowest reliable Weber fraction for the best subject of this experiment was .19, and was found between brightnesses of 9.84×10^{-2} and 323 millilamberts. This value indicates the superiority of the cat in visual brightness discriminations over the rat and raccoon, approximates those found in prior studies on the dog, and is inferior to similar limens which have been established with the monkey.

4. The absolute brightness threshold was found by the present method to be in the vicinity of 1.32×10^{-7} millilamberts. The cat is thus capable of responding to a lower stimulus luminosity than the human individual. This probably does not mean that the cat has a greater retinal sensitivity but rather that the optical speed of the cat's eye is greater. Measured in terms of photons, it is likely that the retinal sensitivity of these two organisms would be roughly equivalent.

5. When the Weber ratios are expressed as a function of changing illumination, the resulting curve for the cat is seen to resemble those established for organisms of a more primitive level in phylogeny (Mya, Drosophila, and honey bee) and the curve for man. As in these organisms, visual intensity discrimination in the cat is poor at extremely low brightness levels. There is a rapid increase in sensitivity as the luminosity rises above the absolute threshold, the Weber fraction approaches a relatively constant value at median illumina-

tions, and there is some indication that sensitivity falls off at very high brightnesses. Irregularity in the data led to speculation about the possibility of a rod-cone shift in the brightness discrimination function in the cat.

6. The variability in thresholds found upon retesting an animal at a given brightness level leads the writer to suggest that any study of infra-primate vertebrate vision which utilizes a discrimination method should take some measure of the reliability of obtained thresholds. Little or no mention of this factor has been made in the published literature.

7. Cats transfer the brightness discrimination habit from one luminosity level to another very readily. The percentage of correct responses made by any of the three animals here used on the first day of an experimental session in which a new brightness level was introduced was almost always above the criterion of discrimination (76 per cent correct in 25 trials). This indicates that the animals were responding to relative rather than absolute brightnesses.

8. The cat data were fitted by the method of least squares to the limiting cases of Hecht's family of theoretical curves. The cat data are not sufficiently precise as to permit the selection of any single one of this family of curves for their description. That is, the plot of the experimentally-obtained points deviates to some degree from the smooth theoretical curves. Nevertheless, the general contour of the empirical curve appears to follow that of the theoretical one, thereby suggesting that Hecht's theoretical photochemical description of the brightness discrimination response has more widespread applicability to the phylogenetic series than has hitherto been shown experimentally.

9. The final conclusion is made that the method and results obtained up to the present time are sufficient to quantitatively describe the brightness discrimination performance of the normal cat. Using these results as a criterion, the present technique offers a means of accurately measuring the effects of operative destruction of the central visual centers. By comparing the curve obtained in the normal cat to that found in animals which have undergone experimental lesions in the central optic tracts, we may hope to obtain a partial answer to the question of the relative importance of the peripheral and central mechanisms which govern this visual function.

References

1. Aubert, H. Physiologie der Netzhaut. Breslau: Morgenstern, 1865. Pp. 394.
2. Bills, M. A., & Maukin, O. Preliminary study of effect of methyl alcohol fumes on brightness discrimination in the white rat. *J. Comp. Psychol.*, 1921, **1**, 495-504.
3. Blanchard, J. The brightness sensibility of the retina. *Phys. Rev.*, 1918, **11**, 81-99.
4. Bridgman, C. S. The absolute brightness threshold and the scotopic visibility curve of the cat's eye. Ph.D. thesis, Univ. of Rochester, 1938. Pp. 110.
5. Chievitz, J. H. Ueber dan Vorkommen der Area centralis in den vier höheren Wirbeltierklassen. *Arch. Anat. Physiol.* (Anat. Abt.), 1891, 311-334.
6. Craik, K. J. W. The effect of adaptation on differential brightness discrimination. *J. Physiol.*, 1938, **92**, 406-421.
7. Crawford, M. P. Brightness discrimination in rhesus monkey. *Genet. Psychol. Monog.*, 1935, **17**, 72-162.
8. Crozier, W. J., & Cole, W. The phototropic excitation of Limax. *J. Gen. Physiol.*, 1929, **12**, 669-674.
9. Folger, H. T. A quantitative study of reactions to light of Amoeba. *J. Exper. Zool.*, 1924, **41**, 261-291.
10. ———. Reactions to light by Mya arenaria in relation to the Bunsen-Roscoe law. *Anat. Rec.*, 1926, **34**, 115.
11. Frolov, G. P. Reaction of the central nervous system to changes in intensity of illumination. *Proc. Petrograd. Soc. Nat. Sci.*, 1918, **69**, 3-19. (Reported by Razran and Warden, 45).
12. Ghiselli, E. E. The superior colliculus in vision. *J. Comp. Neurol.*, 1937, **67**, 451-467.
13. Hecht, S. Sensory equilibrium and dark adaptation in Mya arenaria. *J. Gen. Physiol.*, 1919, **1**, 545-558.
14. ———. The nature of the latent period in the photic response of Mya arenaria. *J. Gen. Physiol.*, 1919, **1**, 657-666.
15. ———. The photochemical nature of the photosensory process. *J. Gen. Physiol.*, 1920, **2**, 229-246.
16. ———. The visual discrimination of intensity and the Weber-Fechner law. *J. Gen. Physiol.*, 1924, **7**, 235-267.
17. ———. Vision: II. The nature of the photoreceptor process. In *A Handbook of General Experimental Psychology*. C. Murchison, ed. Worcester: Clark Univ. Press, 1934. (Pp. 704-828.)
18. ———. A theoretical basis for intensity discrimination in vision. *Proc. Nat. Acad. Sci.*, 1934, **20**, 644-655.
19. ———. Theory of visual intensity discrimination. *J. Gen. Physiol.*, 1935, **18**, 767-789.
20. ———. Intensity discrimination. *Cold Spr. Har. Sympos. Quant. Biol.*, 1935, **3**, 230-236.
21. ———. Intensity discrimination and its relation to the adaptation of the eye. *J. Physiol.*, 1936, **86**, 15-21.

22. ———. Rods, cones, and the chemical basis of vision. *Physiol. Rev.*, 1937, **17**, 239-290.

23. HECHT, S., & WALD, G. The visual acuity and intensity discrimination of Drosophila. *J. Gen. Physiol.*, 1934, **17**, 517-547.

24. HELMHOLTZ, H. Treatise on Physiological Optics: II. The sensations of vision. (J. P. C. Southall, ed.) Ithaca, N. Y.: Opt. Soc. Amer., 1924. Pp. 480.

25. KAPPAUF, W. E. The relation between brightness and critical frequency for flicker discrimination in the cat. Ph.D. Thesis, Univ. of Rochester, 1937. Pp. 94.

26. KÖNIG, A., & BRODHUN, E. Experimentelle Untersuchungen über die psychophysische Fundamentalformel in Bezug auf den Gesichtsinn. *Suzber d. Akad. d. Wiss.*, 1889, 641.

27. KRAUSE, W. Die Retina. Die Retina der Säuger. Inter. *Monatsschr. f. Anat. u. Physiol.*, 1895, **12**, 46-186.

28. LASHLEY, K. S. The mechanism of vision: II. The influence of cerebral lesions upon the threshold of discrimination of brightness. *J. Genet. Psychol.*, 1930, **37**, 461-480.

29. ———. The mechanism of vision: XII. Nervous structures concerned in the acquisition and retention of habits based on reactions to light. *Comp. Psychol. Monog.*, 1935, **11**, 43-79.

30. LOEB, J., & EWALD, W. F. Ueber die Gültigkeit des Bunsen-Roscoe'schen Gesetzes für die heliotropische Erscheinung bei Tieren. *Centr. Physiol.*, 1914, **27**, 1165-1168.

31. LOWRY, E. M. The photometric sensibility of the eye and the precision of photometric observations. *J. Opt. Soc. Amer.*, 1931, **21**, 132-136.

32. MARQUIS, D. G. Effects of the removal of the visual cortex in mammals, with observations on the retention of light discrimination in dogs. *Assoc. Res. Ner. Men. Dis.*, 1934, **13**, 558-592.

33. ———. Phylogenetic interpretation of the functions of the visual cortex. *Arch. Neurol. & Psychiat.*, 1935, **33**, 807-815.

34. MEAD, L. C. The curve of visual intensity discrimination in the cat. *Psychol. Bull.*, 1937, **34**, 517 (Abstract).

35. ———. The curve of visual intensity discrimination in the cat before and after removal of the striate area of the cortex. Ph.D. Thesis, Univ. of Rochester, 1939. Pp. 242.

36. ———. Thresholds of visual intensity discrimination in phylogeny. *Amer. J. Psychol.*, 1939, **52**, 465-467.

37. ———. Visual brightness discrimination in the cat as a function of illumination. *Psychol. Bull.*, 1940, **37**, 501-502. (Abstract).

38. MOODY, P. A. Brightness vision in the deermouse—Peromuscus maniculatus gracilis. *J. Exper. Zool.*, 1929, **52**, 367-405.

39. MUNN, N. L. Pattern and brightness discrimination in raccoons. *J. Genet. Psychol.*, 1930, **37**, 3-34.

40. MUNN, N. L., & SLATER, J. E. A note on brightness vision in the white rat. *J. Comp. Psychol.*, 1932, **13**, 273-277.

41. MURR, E. Ueber den Helligkeitsinn der Hauskatze und die Bedeutung des Tapetum Lucidum. *Zool. Anz.*, 1928, **3**, 254-265.

42. ———. Zur besserem Unterscheidung von Stäbchen und Zapfen in der Retina der Wirbeltiere. *Zsch. f. Zellforsch.*, 1930, 10, 386-393.

43. PIERON, H. La loi de l'excitation lumineuse chez Mya arenaria. *Bull. Soc. Biol.*, 1925, **93**, 1-35.

44. RAWDON-SMITH, A. F. Theories of Sensation. Cambridge: Cambridge Univ. Press, 1938. Pp. 137.

45. RAZRAN, H. S., & WARDEN, C. J. Sensory capacity of dogs as measured by the CR method (Russian schools). *Psychol. Bull.*, 1929, **26**, 202-222.

46. REEVES, C. D. Discrimination of light of different wave lengths by fish. *Behav. Monog.*, 1919, **4**, 1-106.

47. SCHULTZE, M. Zur Anatomie und Physiologie der Retina. *Arch. f. mik. Anat.*, 1866, **2**, 175-286.

48. SLATER, J. E. Brightness vision in the albino rat. *Univ. Pitt. Bull.*, 1931, **7**, 394-395.

49. SMITH, J. R. Spatial and binocular effects in human intensity discrimination. *J. Gen. Psychol.*, 1936, **14**, 318-345.

50. SMITH, K. U. Apparatus for the study of discrimination in the higher vertebrates. *J. Genet. Psychol.*, 1933, **43**, 462-466.

51. ———. An investigation of postoperative deficiencies in visual intensity discrimination following extirpation of the striate cortex in the cat. *Psychol. Bull.*, 1935, **32**, 746-747. (Abstract).

52. ———. The postoperative effects of removal of the occipital cortex upon visual intensity discrimination in the cat. *Amer. J. Physiol.*, 1936, **116**, 145-146. (Abstract).

53. ———. Visual discrimination in the cat: V. The postoperative effects of removal of the striate areas upon intensity discrimination. *J. Genet. Psychol.*, 1937, **51**, 329-370.

54. SMITH, K. U., & KAPPAUF, W. E. Forms of learning in animals. Film, shown at A. P. A. meetings, State College, Pa., 1940.

55. SPENCER, L. T. Central inhibition in the albino rat. *J. Comp. Psychol.*, 1923, **3**, 389-408.

56. STEINHARDT, J. Intensity discrimination in the human eye: I. The relation of $\frac{\Delta I}{I}$ to intensity. *J. Gen. Physiol.*, 1936, **20**, 185-209.

57. STONE, C. P. Notes on light discrimination in the dog. *J. Comp. Psychol.*, 1921, **1**, 413-431.

58. TUGMAN, E. F. Light discrimination in the English sparrow. *J. Anim. Behav.*, 1914, **4**, 85-109.

59. WOLF, E. The visual intensity discrimination of the honey bee. *J. Gen. Physiol.*, 1933, **16**, 407-422.

60. ———. On the relation between measurements of intensity discrimination and of visual acuity in the honey bee. *J. Gen. Physiol.*, 1933, **16**, 773-786.

61. ———. An analysis of the visual capacity of the bee's eye. *Cold Spr. Har. Sympos. Quant. Biol.*, 1935, **3**, 255-260.

62. YERKES, R. M. The reaction time of Gonionemus murbachii to electric and photic stimuli. *Biol. Bull.*, 1904, **6**, 84-95.

63. ———. The Dancing Mouse. New York: Macmillan, 1907. (Pp. 91-132.)

64. YERKES, R. M., & WATSON, J. B. Methods of studying vision in animals. *Behav. Monog.*, 1911, 1, No. 2, 1-90.

65. ZURN, J. Vergleichend histologische Untersuchungen über die Retina und die Area centralis retinae der Haussäugethiere. *Arch. Anat. u. Physiol.*, 1902, Suppl., 99-144.

Department of Psychology
Tufts College
Medford, Massachusetts

The Journal of Genetic Psychology, 1942, 60, 259-269.

SIMPLE AND COMPLEX BASES FOR THE SOLUTION OF VARIOUS TESTS IN THE CARMICHAEL ELEVATED-MAZE SITUATION*[1]

New York University College of Medicine

CHARLOTTE MARKER AND MARTIN DOLGIN

In this journal (1938, Vol. 52), Dr. Leonard Carmichael (1) published some results for a series of experiments with rats in a special elevated-maze problem (Figure 1). The animals were first

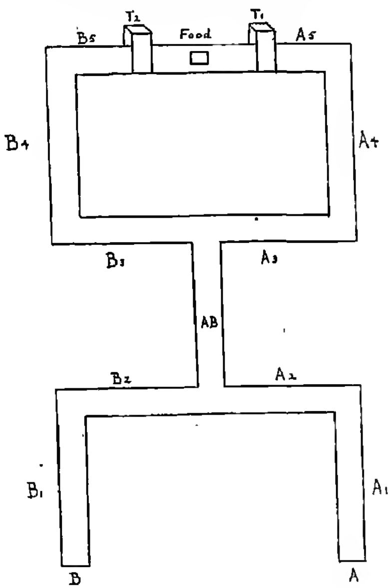

FIGURE 1
DIAGRAM OF MAZE

*Received in the Editorial Office on November 20, 1940.

[1]These experiments were performed in the Animal Psychology laboratory at Washington Square College, New York University, under the guidance and supervision of Professor T. C. Schneirla. The writers wish to express their appreciation to Dr. Schneirla for his invaluable aid in the conduct of the experiment and the preparation of the manuscript.

required to learn a regular alternation habit consisting of A - $A1$ - $A2$ AB - $B3$ - $B4$ - $B5$ and B - $B1$ - $B2$ - AB - $A3$ - $A4$ - $A5$. Then they were given an irregular (chance) alternation to learn, in the same situation. After this, the animals were required to learn a regular alternation habit consisting of A - $A1$ - $A2$ - AB - $A3$ - $A4$ - $A5$ and B - $B1$ - $B2$ - AB - $B3$ - $B4$ - $B5$. Performance in this latter problem was found to be conspicuously poorer than in the initial one. Carmichael's conclusion was that learning, in the first situation, caused inhibitory neural changes which modified the animals' capacity for subsequent learning. We question this as a sufficient explanation.

Although Carmichael's conclusions may prove to be correct upon further experimentation, they are not validated by the research which he completed. That research did not cover certain explanations which may be entertained as alternatives and as additions to the one offered by the experimenter. First, the possible influence of centrifugal swing was not explored. If this complicated factor is not excluded, a run from A to T_2 may be mechanically easier than a run from A to T_1. Thus, for one thing, it is possible that Carmichael's results were due simply to a fundamental difference in difficulty between the two types of runs.

Second, Carmichael apparently did not consider the possibility that the two opposing situations which he used as his testing ground may be fundamentally different for the animal. On the face of the matter, by utilizing alternation between the two starting points and food box entrances as the crux of his experiment, Carmichael put into effect two opposed possibilities for learning. The first (and the one which he evidently believed to have prevailed) is the establishment of a relationship between the starting point and the entrance to the food box. The other, much simpler to learn, is the establishment of a simple alternation habit based upon the tunnel entrances to the food box. But from the report, we cannot choose between these processes as alternatives or as a possible combination. A clue might be obtained from the results of the irregular-alternation (or chance alternation) training, but Carmichael finds nothing of significance here.

The question of what the animal must learn to solve the problem is highly important and its answer must be known if we are to explain adequately the results of the experiment. To obtain a better understanding of the rat's performance in the maze situation, we have repeated the experiment and have added some tests

designed to reveal any centrifugal-swing effect and any influence of the chance-alternation training. In addition, we have changed the procedure in an important respect, concerning the influence of motivation.

The maze pattern was that used by Carmichael. It was constructed by putting together several 26-inch elevated pathway sections. Visual homogeneity was approximated by placing screens around the maze on all sides. The experiment involved (*a*) a repetition of Carmichael's procedures and (*b*) a reversal of Carmichael's procedures. That is, the animals first learned the alternation *A* to *A*5; *B* to *B*5 and then learned *A* to *B*5; *B* to *A*5. The *A* to *B*5; *B* to *A*5 runs we call centrifugal-swing (*CS*) runs, because, theoretically, they are favored by centrifugal swing. The *A* to *A*5; *B* to *B*5 runs are labelled counter-centrifugal-swing (*C-CS*) because, theoretically, *CS* opposes these. This opposition, introduced through a mechanics-of-movement pattern peculiar to the situation, might make this second alternation the more difficult of the two.

Eighteen rats were used in the two experiments. All received preliminary training on a straightaway series of maze units, without choice points. This habituated them to an elevated path, and particularly to opening the doors which gave access to the end-tunnels and, through them, food. One half of the animals in each of the two groups were transferred directly from the original situation to its reverse, i.e., from *CS* runs to *C-CS* and from *C-CS* runs to *CS*. The other half, after they had learned the original problem, were put on the chance alternation schedule. Then, after learning, this situation, they were finally given a new problem (e.g., first *CS*; then chance-alternation of *CS*; finally, *C-CS*).

This involvement of motivation is very important for this problem. In Carmichael's procedure, when the animal first responded incorrectly on a given run, it could then reach food merely by running around to the opposite side. Thus, the animal would receive food on every trial, whether or not its initial response was the correct one. We eliminated this condition by removing the animal from the maze each time it made an incorrect response, and replacing it at the starting point. Similarly, if the animal started at, for example *A*, and ran across to *B*2, it was removed and replaced at the starting point. This prevented the fixation of an "incorrect" starting point-food box relationship.

The criterion for learning was six consecutive errorless runs.

Results

In Table 1 we present the results for Group I. It will be seen

TABLE 1
TRIALS REQUIRED TO LEARN THE SERIES: *CS* REGULAR ALTERNATION PROBLEM, *CS* IN CHANCE-ALTERNATION PROBLEM (INTERPOSED FOR CERTAIN ANIMALS), AND *C-CS* REGULAR ALTERNATION PROBLEM

Rat number	*CS* alternation	*Trials to Learn* *CS* in chance alternation	*C-CS* alternation	Ratio of *CS* : *C-CS*
1	4	81	18	1 : 4.5
2	7	17	12	1 : 1.7
3	13	—*	17	1 : 1.3
4	3	18	5	1 : 1.67
5	5	—	28	1 : 5.6
6	23	—	3	1 : .16
7	14	—	7	1 : .15
8	9	96**		
Mean no. of trials required	9.7 (8 cases)	35.3 (3 cases)	12.8 (7 cases)	

*— signifies that no chance-alternation training was given to the animal.
**The experiment was stopped after a total of 105 trials, although the animal had not yet learned the *CS*-chance-alternation pattern.

that five of the first seven animals required more trials to master the *C-CS* problem, although this test *followed* the *CS* tests. (In three cases, it also followed the *CS*-chance-alternation tests.) The number of trials to learn the chance-alternation was in each case greater than the number required for the initial learning. The ratios of trials of *CS* : *C-CS* required by these animals ranged from 1:0.15 for No. 7 to 1:5.6 for No. 5. The mean number of required trials was 9.7 for *CS* learning, 12.8 for *C-CS* learning, and 35.3 for *CS* chance-alternation learning. In Table 2 we state the number of trials required for successive tests by Group II. These animals were given the *C-CS* test first of all. Six of the nine animals required *more* trials to learn the *C-CS* situation than to learn the *CS* situation, which was presented last to the group. The ratios of required trials (*CS*=1.0) ranged between 0.2 for No. 15 and 26.0 for No. 17. It may also be noted that No. 7 required only one trial to master *CS*, 26 to master *C-CS*. The mean number of trials required was 17.3 for *C-CS*, 14.2 for *CS*, and 25.0 for *C-CS* chance-alternation learning. No. 16 required 68 trials to reach the criterion for learn-

TABLE 2

Trials Required to Learn the Series: *C-CS*-Alternation Problem, *C-CS* in Chance-Alternation Problem (Interposed for Certain Animals), and *CS*-Alternation Problem

Rat number	*C-CS* alternation	*Trials to Learn* *C-CS* in chance alternation	*CS* alternation	Ratio of *CS* : *C-CS*
9	27	—*	16	1 : 1.7
10	14	4	9	1 : 1.6
11	27	—	4	1 : 6.7
12	4	18	9	1 : 0.44
13	15	—	8	1 : 1.9
14	18	—	16	1 : 1.1
15	12	10	52	1 : 0.2
16	13	68	14	1 : 0.9
17	26	—	1	1 : 26.0
Mean no. of trials required.	17.3	25.0	14.2	

*— signifies that no chance-alternation training was given to the animal.

ing of the chance-alternation pattern, but both Nos. 10 and 15 required a smaller number of trials to master this problem than the initial one.

Discussion of Results

We have stressed the fact that centrifugal swing (Schneirla, 2) may be of direct or indirect importance as an influence upon the animal's adjustment to the Carmichael situation. Its rôle in the rat's alley-maze response has been conclusively demonstrated by Witkin and Schneirla (3) and by Witkin and Granich (4). The former experimenters found that when a choice-point was preceded by an alley-pattern such as our *CS*-arrangement (i.e., with opposed turns), the outside alternative was favored over the other turn in 59 per cent of the cases; when the choice-point is preceded by a *U*-series of alleys (which we do not employ) the outside alternative was favored in 86 per cent of the cases.

Although the involvement of centrifugal swing in the rat's elevated-maze responses has not been studied as yet, our observations and results suggest that it is involved. But in any maze study it is premature to search for special factors determining relative difficulty of parts or solution of the whole, before this fundamental factor has been taken into account.

If centrifugal swing is a factor in the Carmichael situation, then it would operate in favor of the A-T_2 and B-T_1 runs (previously referred to by us as *CS* runs), and there should be some indication that these are more easily mastered than the routes referred to as *C-CS*. The trend of the results reported in Tables 1 and 2 definitely shows this to be the case for the greater part of the animals; five out of seven when the *CS* run was presented first, six out of nine when the *CS* run was presented after the *C-CS* runs were learned. The latter set of results, presented in Table 2, is particularly striking since it shows that the earlier learning of the *C-CS* situation did not cause animals Nos. 9, 11, 13, 14, and 17 to have even more difficulty in mastering the *CS* situation than they had with the situations first presented (i.e., *C-CS*). Thus, whether the *CS* situation was presented first of all, or later in the series, it proved to be less difficult than the *C-CS* alternative for most of the animals.

But centrifugal swing, although it appears to be a factor of basic effect here, is definitely not the only factor, as Carmichael's (1) findings and other features of our results indicate. Let us examine the required performance in some detail. A regular alternation, such as was demanded of the animals in most of the tests of this experiment, is essentially a lower-level type of performance, since it requires the animal merely to maintain a regular series of position habits. It must be borne in mind that the Carmichael situation contains two spatial alternatives in each of its principal sections, i.e., two starting points and two end points. It is of great importance to determine whether the animal may master any test set for it simply in terms of an alternation of responses to the critical junction in the pathway joining these points, or must go beyond this in its learning. This latter possibility for solution, which Carmichael adopts in his conclusion, is that the animal may establish a starting point-end point relationship. This, if and when it comes into play, must be adjudged a higher-level type of adjustment.

We are faced with the necessity of determining whether only one of these possible adjustments figures in the animal's mastery of a given problem in this situation, or whether both of them are involved. Carmichael's report does not furnish a clear basis for deciding. Our own tests were planned with a view to solving this difficulty.

If the animal masters a problem merely in terms of an alternation of response based upon *limited spatial cues,* then traces of a positive transfer of this alternation habit should be evidenced in a subsequent

learning test. For instance, if the animal has first learned the *C-CS* alternation and is then transferred to the *CS* alternation test, on this basis we should expect his performance to be correct from the beginning in the new situation. That is, it should be correct if the transferred response is merely a habit of alternating the direction of turning at the central choice-point, rather than a more highly organized response. Other complications should arise if there has been established a relationship between starting point and tunnel-foodbox during the previous test.

Our results indicate that in most of the cases this *relational organization* was involved to some extent, since all of our animals, with the exception of No. 17, required at least a few trials to learn the new problem. But there are indications that the simpler alternation habit is more persistently and more frequently effective in the solution of the regular-alternation problems than is the relational habit. This conclusion is encouraged by the fact that, when animals were presented with a *CS*-alternation test directly after mastery of the *C-CS* alternation problem, performance in each case was better on the second problem than on the first.

There would appear to be a conflict between this alternation-response hypothesis and the involvement of centrifugal swing, but it would be a mistake to consider these factors mutually exclusive. It is our contention that the two effects furnish the different components of one solution-pattern—that their influence is additive.

This view of the matter is borne out by the results for animals Nos. 3, 5, 6, and 7. These animals were initially presented with the *CS* alternation problem and later with the *C-CS* alternation problem. It is to be seen that for these rats, in the subsequent *C-CS* alternation test, there is a definite possibility for conflict between centrifugal swing and the positive transfer of an alternation habit. That is, a positive transfer of the simple alternation would permit a facile solution of the somewhat different alternation problem, whereas *CS* would oppose this easy solution. If the trials required for solution of the second problem show an increase over the number required for the first, then we might conclude that *CS* had dominated. But the outcome varies from animal to animal. In the cases of animals Nos. 3 and 5, centrifugal swing was the dominant factor; in the cases of animals Nos. 6 and 7, a transfer of alternation response apparently carried the day. Furthermore, the conspicuous agreement in the transfer results for animals Nos. 9, 11, 13, 14, and 17 is

easily explained if we assume that these two variables operated there in the same direction, that is, both favoring a reduction of trials in the solution of the second problem.

With reference to the above, we wish to stress at this point the significance of some data concerning the importance of CS in the experiment. The mean trials for *original learning* when a CS alternation had to be mastered was found to be 9.75 trials. In contrast, the mean for *original learning* of the C-CS alternation was 17.3 trials. Although the number of animals precludes any statistical check on the reliability of this difference, we do feel justified in accepting the outcome as a strong indication that there existed a difference in difficulty between the two types of runs.

The significance of our chance-alternation tests may now be considered. It is our contention that this situation cannot be learned as a simple alternation habit but that, in order to learn the problem, the animal must establish a relationship between the starting points and the tunnels. If this is the case, the chance-alternation learning should require a greater number of trials for solution than the regular alternation pattern. Furthermore, from the above hypothesis, we should expect that, if a chance-alternation test were interposed between the initial and final learning, it would have a definite effect on the latter. This would be expected if the chance-alternation learning demanded a type of learning not strongly involved in the initial test for which an alternation-response might be sufficient (in other words, if the chance-alternation introduces a new problem). Our results show that the chance-alternation pattern offers no significant elements which can be transferred positively to the final problem. The evident reason is that relationships are established in the chance-alternation procedure which are inimical to the solution of the last regular-alternation problem.

We have previously offered reasons for concluding that a simple-alternation response dominated the learning in the first of the series of problems presented to the animals. This does not preclude the possibility, however, that the animals may have "adopted" the more complex starting point-tunnel relationship as the basis for learning this problem. The latter possibility would be expected to arise with relative infrequency. However, in the event of its occurrence, we would expect the number of trials required to learn the chance-alternation pattern (which follows the initial learning) to decrease considerably. The fact is that with one exception all of the animals

presented with the chance-alternation problem as a second test showed the expected increase in required number of trials. Circumstances suggest that the single exception, No. 10, may have progressed beyond the described lower-level response.

The *CS*:*C-CS* ratios given in Tables 1 and 2 disclose the effect of the chance-alternation test as an influence upon subsequent learning. In the case of Group I, both the effects of centrifugal swing and the chance-alternation test might offer some opposition to the learning of the following *C-CS* alternation problem. Table 1 shows that the number of trials required to solve the latter test is (in the case of every animal given the chance-alternation test) greater than the number required to solve the original problem. On the other hand, among the four animals not given the chance-alternation test, two solved the *C-CS* pattern in fewer trials than were required to solve the original problem. For Group II (Table 2) for which centrifugal swing favored solution of the final problem, all of the animals who had chance-alternation (except No. 10) showed an increase in the number of trials required for the *CS*-alternation. Evidently a conflict in response is present here, but the effect of the chance-alternation training dominates over the effect of *CS* for all animals except No. 10. On the other hand, those animals given the *CS*-alternation test immediately after learning the *C-CS* alternation problem required fewer trials to solve the *CS* test than the original test. This suggests an absence of conflict in that case.

This evidence shows that learning the chance-alternation test conflicts with and hinders the solution of a subsequent learning problem such as Carmichael gave his animals. Our explanation of the difficulty is that none of the learned elements from the chance-alternation can be transferred positively to the new problem situation. The organization formed in solving the chance-alternation problem is diametrically opposed to the one which is required for the solution of the following problem. This is borne out by our finding that the number of trials required for the solution of the problem which followed a chance-alternation test was greater than the number required for solution of the initial test, even when the last test was ordinarily an easy one (from the point of view of centrifugal swing facilitation).

The fact that the mastery of a chance-alternation problem proved to be more difficult than either of the problems which followed is understandable from our explanation. The last problems were tests

which we have found to require merely a simple alternation habit, whereas the chance-alternation problem required the establishment of a starting-point—end-point relationship.

Carmichael's statement that learning in one situation causes neural changes which modify the animal's capacity for subsequent learning is undeniably true as an isolated statement. But the important question is concerned with the different factors in learned adjustments which underlie such modifications.

We have repeated and extended Carmichael's test because his results did not bring out the nature of the forces involved here. Our results show that what is learned in this situation, thus what is available for transfer, depends upon the factors selected by the procedure. For instance, we consider our practice of giving food only at the foodplace which was "correct" according to the nature of the test, of great importance for the outcome. According to the experimenter's manipulation of the routine, the animal learns in this situation either a lower-order alternation habit or a higher-order organization involving the relationship of two spatially separated locales, these two factors being further complicated by centrifugal swing effects.

Summary and Conclusions

Carmichael's experiment (1) has been repeated and extended in order (*a*) to test the involvement of centrifugal swing as a factor in the mastery of various problems in his situation; (*b*) to discover the manner in which the chance-alternation problem is solved; and (*c*) to investigate the nature of the animal's learning under different conditions, in the light of its transferred effect upon other subsequent problems.

Though the problem involves alternation responses on a simple maze as was used here, two items combined through learning account for the animal's solution. The first of these is centrifugal swing, which operates differently according to the problem setting, and the second is the habit of alternating responses to a choice-point. We have concluded that although this pattern centers about centrifugal swing, the alternation-response is the dominant element in it. These two factors operate simultaneously and the sum of their effects depends on the experimental conditions.

The difficulty of this problem varies according to how the routine introduces centrifugal swing. The test which required a turn against

centrifugal swing representatively proved to be more difficult than the test in which the schedule favored a centrifugal swing turn.

The chance-alternation problem requires a more complex learning than do the regular alternation tests, because it cannot be solved on the basis of simple alternation of position habits. This test demands the organization of a relationship between starting-point and ending of the situation. The evidence suggests that this higher-level organization interferes with the subsequent learning of a regular alternation problem.

References

1. Carmichael, L. Learning which modifies an animal's subsequent capacity for learning. *J. Genet. Psychol.*, 1938, **52**, 152-163,
2. Schneirla, T. C. Some comparative psychology. *J. Comp. Psychol.*, 1933, **16**, 307-315.
3. Witkin, H. A. & Schneirla, T. C. Initial maze behavior as a function of design. *J. Comp. Psychol.*, 1937, **23**, 275-304.
4. Witkin, H. A. & Granich, L. An application of some principles of maze-mechanics in the planning of a serviceable maze. *J. Comp. Psychol.*, 1937, **24**, 523-545.

New York University College of Medicine
477 First Avenue
New York City

The Journal of Genetic Psychology, 1942, 60, 271-305.

A STUDY OF SOCIAL STATUS ON THE SECOND GRADE LEVEL*

North Texas State Teachers College

MERL E. BONNEY[1]

A. PROBLEM AND METHOD

The emphasis on socialization as an important part of education has led to much discussion and some research in recent years. Such topics as leadership, factors associated with social success and failure, reasons for liking and disliking people, mutual friendships, and techniques for stimulating social efficiency have been given a rather large amount of attention in educational and psychological literature. Having become interested in these topics, the writer initiated a study of various phases of social success among primary grade children in Denton, Texas, in the fall of 1938. Denton is a town of approximately 12,000 population and is the location of two state colleges, The North Texas State Teachers College and the Texas State College for Women. The data for the study herein reported were gathered from three schools—the Demonstration School, which is conducted as a part of the North Texas State Teachers College, and two public schools known as the Sam Houston and Robert E. Lee Schools. Throughout this report these schools will be designated as Schools *A*, *B*, and *C*.

In order to avoid repeating the number of cases in each school numerous times it can be stated at this point that the population of School *A* fluctuated between 23 and 28 pupils during the course of this study; that of School *B* between 30 and 38, that of School *C* between 24 and 37—with the most typical figure being near 30.

The main purpose of this investigation is to determine the factors related to social success and to devise means of raising the social

*Received in the Editorial Office on November 23, 1940.

[1]The writer wishes to express his appreciation to the school officials and teachers of Denton who have given such excellent coöperation in this study. These include: Superintendent R. C. Patterson, Dr. J. C. Matthews, Mr. J. L. Yarbrough, Mr. C. J. Walker, Miss Bess Tilford, Miss Epsie Young, Mrs. Bess Wilkins, Mrs. Una Gibbons, Miss Roberta Rogers, Mrs. Frances W. Wiley, Miss Inez Odell, and Mrs. C. C. Williams. Appreciation is also due the coöperating Denton physicians: Dr. W. H. Hawley, Dr. Charles Saunders, Jr., and Dr. Jack Woodward.

acceptance of those who are below average. This report, however, does not deal with the second phase of this problem. The study was started on first-graders during the school year 1938-39 but nearly all the data in this report were obtained from these children while they were in the second grade. During the first year most of the time was spent using various time-sampling devices of social behavior with the primary object of becoming personally acquainted with the children to be studied.

The terms "social status," "social recognition," "social success," and "popularity" are used interchangeably in this discussion. This status, recognition, success, or popularity was determined by pupil choices.[2]

In School *A* there were five choosing situations as follows: October, stating preferences of children with whom to have picture taken; November, repeat of the same situation; December, listing children to whom Christmas presents were to be given; April, stating choices of children they would prefer to go home with after school; May, selecting one preferred as a partner on the bus trip taken by the group. In the first two situations the children made a first, second, and third choice, but in the remaining three situations they were instructed to list as many as they desired without any limit.

In School *B* there were six choosing situations as follows: October, selecting children with whom to have pictures taken; January, choosing a child to sit by at the beginning of the second semester; February, listing pupils to whom valentines were to be given; March, selecting a partner for the group trip to the zoo; May, listing preferences of children to go home with after school. In October the choosing was limited to first, second, and third choices. In all the others the number selected was unlimited. In addition to the classroom situations, pupil choices were also obtained in School *B* in 22 playground games from October to January.

In School *C* there were eight choosing situations as follows: October, selecting others in the room with whom to have picture taken; November, listing names of all those to whom presents were to be given at Christmas party; January 8, choosing partners for class trip on a train; January 21, naming children they would prefer to

[2]The writer is indebted to J. L. Moreno's *Who Shall Survive* and his *Sociometry—A Journal of Interpersonal Relations* for stimulation in seeing the possibilities of sociometric studies.

sit by at beginning of second semester; February, listing names of children to whom valentines were to be given; March, electing a child to be class librarian; April, choosing a child to hold a fossil while a group picture was taken (the fossil had played an important part in a class project); May, listing preferences of partners for marching to the Teachers College to see a flower exhibit. Except in the case of the October choosing there was no limit to the number of selections that could be made.

The scoring technique was as follows: first choice 5 points, second 4 points, third 3 points, fourth 2 points, fifth 1 point, and all choices thereafter 1 point. This is admittedly a crude device for reducing to quantitative form such a subtle thing as differing degrees of social acceptance. However, when six or eight choosing situations are held during a school year the final score obtained for each child is certainly not far from designating his actual status in the group. Teacher judgment on the whole supported the pupil choices. Another small check was had in School *C* in which a popularity contest was sponsored by school officials for the purpose of an assembly program. In the second grade the two pupils elected corresponded exactly to the two top pupils on the basis of the status scores for the entire school year.

The composite score for each child was determined by first converting his raw score in each choosing situation into a per cent and then averaging his per cent scores in all of the selections made throughout the year. This made his total average per cent score. In case a child was absent when choices were made it was sometimes possible to have him make his selections when he returned. This was not always feasible, however, so some children in each school did not have scores for all the choosing situations. In order not to lose these cases from the total average per cent scores each child who was absent on a particular choosing occasion was given a score equal to the average score he received when he was present. In a few instances the teacher wrote the names of all the children on the blackboard and called attention to the fact that a child could be chosen even though he was not present. In such instances no child was counted absent. (Before the children could write they came up to the teacher and whispered their choices to her while she wrote them down.) In determining the list for the full term no child was included who was not present for at least half of the choosing situations. In only six cases in all three schools was this

degree of absence involved (of those included in the lists) and in every case these pupils were extremely low in status scores. Usually the amount of absence was only one or two times. Of course, this method of taking care of absences involves a certain amount of error but it certainly is not large. Nor did these adjustments play a significant rôle in determining status placements since nearly all of those who were most frequently absent from the choosing situations were pupils of very low status so that they received a very small portion of the choices in any case. It should be made clear, too, that the adjustments for absence were made only for the total average per cent scores. In running the correlations between successive choosing situations (reported later) only those cases were used which were present in both situations.

B. Distribution of Status

A number of investigations have afforded data on the distribution of status in various groups when status is measured by sociometric tests. H. H. Jennings (11), in his studies in the New York Training School for Girls, found that when 236 girls were asked to choose a living mate and a working mate 29 per cent were below the mean on three counts, namely, in number of votes received, number of different girls voted for, and number of mutual friends. This means that approximately one third of the girls were in a decidedly inferior position in respect to friendship. Jennings found that 6 per cent of the girls were isolates since they received no choices at all. Moreno (12) reports from a study of 505 girls in the New York Training School for Girls that after a number of successive choices 15 per cent of the girls remained unchosen. Dimock (4) had 118 boys in a summer camp choose 10 other boys for tent-mates. He found that 25 per cent of the boys received no first or second choices and that 14 per cent were chosen only in last or next to last places.

The distribution of status in the three Denton schools is shown in Table 1.

It is evident from the above table that the distribution of social recognition in the groups studied does not follow a normal distribution curve. Instead there is a high concentration in a selected group at the top. This point is further emphasized by noting the difference between the total score votes of the highest pupil in each group and the ones at the bottom. In School *A* the highest child

TABLE 1
DISTRIBUTION OF STATUS SCORES IN THE SECOND GRADES OF THREE DENTON SCHOOLS

School	Per cent of total average per cent scores of entire group held by upper 20 per cent of pupils	Per cent of total average per cent scores of entire group held by lower 20 per cent of pupils
A	58	8
B	48	5
C	54	3

received a total score equal to that of 11 pupils (48%) at the bottom of the group. In School *B* the highest child received a total score equal to that of the lowest 11 pupils (29%). In School *C* the top child received a total vote equal to the lowest 12 pupils (43%). Although there was no child who failed to receive a single vote throughout the entire year there were a number of isolates in nearly all of the separate choosings. This number varied from 0 to 23 per cent with an average for all three grades of 10 per cent. It was noted that there was a larger number of isolates in those choosing situations in which some ability was involved such as acting as librarian or as nurse or doctor in a school project. The situation in which there were no isolates was in the giving of valentines.

The question may be asked as to whether the children who were most popular received their choices from a wide distribution of children or from a small group of friends? Also, did the most popular pupils vote for as many children as voted for them? These questions are answered in Table 2.

TABLE 2
AVERAGE NUMBER OF DIFFERENT CHILDREN VOTING FOR THE SOCIAL STATUS QUARTILE GROUPS IN THE SECOND GRADES OF THREE DENTON SCHOOLS AND THE RATIO BETWEEN THE NUMBER OF DIFFERENT CHILDREN VOTED FOR BY EACH QUARTILE GROUP AND THE NUMBER VOTING IN TURN FOR THESE CHILDREN

Social status quartiles	School *A* No. of children voting for	Ratio average	School *B* No. of children voting for	Ratio average	School *C* No. of children voting for	Ratio average
4	23	187	21	157	26	241
3	16	101	18	132	18	141
2	13	92	10	86	14	98
1	10	74	7	52	8	77

In respect to the first question above it can be readily seen from Table 2 that it is answered in the affirmative. In each school the average number of different children voting for the quartile groups varies directly with the degree of social acceptance—when the first quartile is understood as being the lowest in status and so on up to the fourth. In School *A* the number of different children voting for the highest quartile group was more than twice as large as the number voting for the lowest group and in Schools *B* and *C* the proportion was 3 to 1 or better. The fact that this kind of relationship held true for each quartile in all three schools definitely shows that on the whole the pupils who were most popular were so because of a wide acceptance throughout their groups rather than because of consistently receiving high votes from a few close friends. This also shows that these groups of children were quite well agreed on the factors which determine social status.

The figures on ratios in Table 2 were obtained by dividing the number of different children which each child voted for throughout the year into the number of different children who voted for this child. The ratios of each quartile group were averaged to make the above figures. However, only those choosing situations were used which involved an unlimited number of choices. It will be clear that when an average quartile ratio is greater than 100 it means that the children in this group had more children voting for them than they voted for in return. Thus the average ratio of 241 for the upper group in School *C* means that for every child which these children choose throughout the year approximately 2½ children choose them. It will be observed that the ratios for all the first and second quartiles are below 100. This shows that the children below average in status in each school are reaching out for more friendship contacts than they are receiving. On the other hand many (but not all) of those who were most popular in each school confined their choices to a few of their best friends. The reason for this situation probably is that those in the upper quartiles feel secure in their status and so do not feel the need of reaching out for more responsive social contacts, whereas most of those below average in status sense their insecurity and are anxious to do anything which offers promise of throwing them into friendly relations with other children. Hence they put down more names in the choosing situations in the hope of thereby pleasing more children,

or of ensuring themselves a partner, as the case might be in the situations where choices were made.

What attitude should be taken toward the fact that studies of various groups including those in Denton show a rather high degree of concentration of status in a few members at the top? Is this a so-called "natural" condition or is it highly susceptible to change? Should it be changed? Would those at the top of the group be better off if social recognition and friendship were more evenly distributed throughout the entire membership? Can those in the lower third or half in status have adequate or normal personality development in spite of the limited positive tele which they arouse in others? In answer to the last question Dimock (4) presents evidence to show that one or two friends is not enough for wholesome personality growth. It is the opinion of the writer that it is highly desirable to stimulate a better distribution of social acceptance throughout a group than sociometric studies show to be the case. In saying this there is no implication that those at the top in status have too much recognition in any absolute sense but only relatively. One of the fortunate things about social acceptance is that it is not a limited quantity which, if it is to be changed, can only be re-divided in a different way. On the contrary it is something which can be created indefinitely without anyone losing what he possesses. If the lower one-third of individuals in status in a particular group were so helped that they raised their social acceptance, say 20 per cent, not only these persons but the entire group would be enriched, including those at the top. The amount of optimism, good humor, mutual appreciation and all that is included under "friendship" would certainly be increased. Would this not be a gain for all? An analogy may be drawn here with the field of health. Are we not all more healthy because of the fact that health is now much better distributed throughout our population than it was one hundred years ago? This has been everyone's gain and nobody's loss. This would seem to be true of all psychological and spiritual goods and is certainly true in some respects at least in the area of material goods.

C. Stability of Status

A question often raised in sociometric studies is in regard to the stability of status. Do those who receive the most votes in one situation continue to rank high in the other choosing situations?

On this point Moreno (12) reports that in the New York Training School for Girls his subjects maintained their original choices for living with each other to the extent of 95 per cent after an interval of three months. In a study made by Joan Criswell (3) in a New York City public elementary school it was found that both first and second choices remained stable over a period of six weeks to the extent of 36 per cent. First choices were most resistant to change—being 69 per cent stable over the six weeks period.

The stability of status scores in the Denton schools is shown in Table 3.

TABLE 3
CORRELATIONS BETWEEN SUCCESSIVE STATUS SCORES IN THREE DENTON SCHOOLS

School *A*	First grade and Oct. 1939	First grade and Sec. grade	Oct. and Nov.	Nov. and Dec.	Dec. and Apr.	Apr. and May	May, 1940 and Oct. 1939
	.74	.82	.73	.83	.64	.76	.77

School *B*	First grade and Oct. 1939	First grade and Sec. grade	Oct. and Jan.	Jan. and Feb.	Feb. and Mar.	Mar. and Apr.	Apr. and May	May, 1940 and Oct. 1939
	.64	.56	.46	.49	.16	.27	.63	.36

School *C*	First grade and Oct. 1939	First grade and Sec. grade	Oct. and Nov.	Nov. and Jan. 8	Jan. 8 and Jan. 21	Jan. 21 and Feb.	Feb. and Mar.	Mar. and Apr.	Apr. and May	May, 1940 and Oct. 1939
	.34	.48	.68	.94	.75	.76	.65	.89	.86	.70

An examination of Table 3 shows that all the correlations are above zero and positive. In Schools *A* and *C* nearly all the coefficients are above +.65. For the second grades alone the average correlations for the three schools are as follows: *A* +.75, *B* +.40, *C* +.78. It will be noted that the relationship between the total first grade standing and the first choosing situation in October, 1939, is rather high except in the case of School *C*. It seems probable that the results for Schools *A* and *B* are the most valid since the data for the first grade were more adequate in these schools than in School *C*. However, the number of cases in each instance was small being 21, 16, and 17 respectively. Approximately this same number of cases was available for the correlations between the

total first grade and total second grade standings. In School *A* this relationship was high and in the other two a fair degree of stability was shown.

It will be observed that the coefficients for successive choosings in School *B* are considerably lower than in the other two. There are two possible explanations of this fact. In the first place the children in this group were more homogeneous than either of the other two second grades in intelligence, social competence, and home backgrounds. Data on the first two of these factors are given in Table 4.

TABLE 4

VARIABILITY OF SECOND GRADE PUPILS OF THREE DENTON SCHOOLS IN *IQ* AND SOCIAL COMPETENCE

School	*IQ*			Social competence		
	Extreme scores	Range	Sigma	Extreme per cent scores	Range	Sigma
A	82-148	66	16.25	.01-.20	19	4.45
B	72-125	53	13.30	.001-.08	7.99	2.15
C	66-137	71	14.80	.002-.14	13.99	3.04

In Table 4 both the range and standard deviation of *IQ*'s show that the pupils in School *B* were more on a par with each other in this respect than were those in the other two schools. Likewise the figures for social competence (total average per cent scores) show much less variability between the upper and lower pupils and a greater homogeneity throughout the distribution. This means that the pupils in School *B* were more equalized in social success.

Unfortunately accurate data on the home background have been obtained on only one of the three schools under study. However, the writer has visited nearly all of the homes involved and has had the impression that those in School *B* are more homogeneous than the other two groups. In School *A* the homes ranged all the way from very poor unpainted country houses to the best homes in Denton. In School *C* the children came from home situations ranging all the way from the worst of shacks to those with incomes between $1,500 and $3,000. There is no doubt but that the range between the upper and lower home groups in School *B* is considerably less than in School *A* and it is the writer's impression that it is also less than in School *C*. More complete data will be obtained on this point during a subsequent follow-up.

The fact that the pupils in School *B* were more homogeneous in intelligence quotients, social competence, and home background made it possible for choices to be better distributed throughout the entire group than was the case in the other schools. Since a few pupils were not so decidedly superior to the others a greater shifting of choices was possible from one choosing situation to another. A larger proportion had a chance to appear in a favorable light in the estimation of others. This would make for lower correlations between successive choices than would be true in a group in which a few pupils were very outstanding in one or more factors important to social status. It may be recalled that in the previous section on distribution of status the upper fifth in School *B* did receive 48 per cent of the total per cent scores cast during the school year. This, of course, is a rather severe concentration in the upper brackets, but it is less than in the other two schools (10 and 6 per cent) and there was more shifting of choices among the pupils in this upper 20 per cent as well as in the rest of the distribution.

Before leaving this point about the greater homogeneity of School *B* it will be appropriate to point out a practical application, namely, that if a child does not possess very outstanding traits he will be much more apt to attain social recognition in a group composed of others of approximately his level of ability to compete for status. This fact has often been observed, but there is little evidence to show that educators have been any where near as much concerned about grouping children for purposes of social success as they have about grouping them for academic success.

The statement was made in a foregoing paragraph that there are two possible explanations of the lower correlations between scores in successive choosing situations in School *B* than in Schools *A* and *C*. Aside from the factor of greater homogeneity in School *B* there is also the factor of greater pupil turn-over. In School *A* there were three new pupils throughout the year and two drop outs. In School *B* there were 13 new pupils and 12 who dropped out. In School *C* there were five new pupils added to the roll while 11 were lost. On the basis of the number present for the first choosing in October those figures represent the following percentages: School *A*, 11 and 7 per cent; School *B*, 36 and 33 per cent; School *C*, 15 and 32 per cent. These data show that there was a marked difference between Schools *B* and *A* in pupil turn-over and a considerable difference between Schools *B* and *C* in number of new pupils added.

It can readily be understood that a group which has a high degree of turn-over would show lower correlations between successive choosing situations than one in which the population is fairly stable. Another factor affecting the correlations from the standpoint of turn-over is the degree of popularity of the pupils gained or lost. In School *A* none of the pupils either gained or lost were even average in recognition. In School *B*, on the other hand, two of the pupils gained became better than average in popularity and two of those lost were in the upper quartile in status. This naturally made for a greater distribution of votes. In School *C* two of those added during the year attained recognition scores in the third quartile whereas only one of those lost had achieved even average status. This explains why in spite of a 32 per cent loss of pupils in School *C* there was a negligible effect on the distribution of status. It may now be concluded on the point under discussion that one of the important factors determining the extent to which recognition scores remain stable throughout a school year is the extent to which the group population remains stable, and more particularly the extent to which there is a turn-over in pupils who are able to attain a high degree of social competence.

There is little doubt but that a teacher plays a significant rôle in determining the stability or instability of status in a school group, but nothing on this point could be decided from the data in this study. In School *A* the same teacher was in charge the full year whereas there was a change of teachers in both the other schools at the end of the first semester. Since, of the three schools, School *C* showed the highest average correlation between scores in successive choices throughout the year this shows that at least in this group other factors were far more important in determining recognition than the change of teachers.

Before closing this discussion on the stability of status one more point remains to be made. This is in respect to the relation between scores made in choosing situations in the classroom and scores made in choosing situations on the playground. As previously stated, data were obtained on this point only in School *B*. The writer had two assistants go down to this school two days a week during the play period. While one assistant conducted the games the other took down the choices made whenever a choosing game such as Farmer in the Dell was played. This was done from the latter part of October to the latter part of January. Altogether, data were

obtained on 22 game-choosing situations, with an average of 15 choices per game. By counting each choice as one point a total score was arrived at for each child. These total scores ranged from 2 to 32 with a mean of 11. After changing these scores into per cents (using 27 cases) a correlation was run with the total average per cents for the whole year. The r proved to be +.62. Also a correlation was run between the playground scores and the classroom choices made on January 21. This was found to be +.76. These results point to the conclusion that on the whole the children who were most popular in the classroom were also most popular on the playground.

D. *IQ*'s and Social Status

Hardy (8), in a study of elementary school children involving choices of friends, found an r of +.37 between *IQ*'s and social recognition. When only the upper 20 per cent in recognition were considered it was found that not one had an *IQ* below 90 and 29 per cent had *IQ*'s of 120 or more. Hsia (10) conducted a study in the upper grades of the elementary school in which a composite social score was determined by combining teacher ratings, pupil choices, and the results obtained on a Sociability Test. Between this composite score and *IQ*'s an r of +.30 was obtained.

McGahan (14) obtained leadership scores on children in the fourth, fifth, sixth, and seventh grades based on six elections held throughout the school year. Between these leadership scores and *IQ*'s he found in the respective grades correlations of +.33, +.46, +.36, +.36.

In the Denton study the *IQ*'s were obtained in both the first and second grades by means of California Mental Maturity Scale. In each grade the tests were administered the latter part of January. To show the relationship between the *IQ*'s and social acceptance the social status scores for the entire year for the three schools were combined into quartiles, *i.e.*, the first quartile in each school was thrown together to make one combined group and so on up through the fourth quartile. After obtaining these combined social status quartiles the average *IQ* of each group was determined. In the first grade (76 cases) the average *IQ*'s for the four groups ran as follows: first quartile 95, second 103, third, 105, and fourth 111. The first quartile is the lowest in status. In the second grade (79 cases) the average *IQ*'s ran: first quartile 92, second 102, third 108,

and fourth 109. These figures show that the relation between *IQ*'s and social acceptance during these two years was not marked. However, in each year the two upper quartiles had higher average *IQ*'s than the two lower groups and the contrasts between the first and fourth quartiles are no doubt significant—being 16 and 17 points respectively in the two years. In the two upper quartiles there were only three children with *IQ*'s under 90 whereas there were 16 lower than this figure among those below average in status.

Pearson Product Moment coefficients were obtained between *IQ*'s and total average per cent scores and social status in the second grade. In School *A* this turned out to be +.02; in School *B*, +.44; and in School *C*, +.11. All these relationships are too low to be significant. A number of additional correlations, six in all, were run between *IQ*'s and scores made in particular choosing situations. These varied from +.04 to +.46. The latter figure was found when the purpose of the choosing situation was to select a librarian. This may have some significance in that there was usually a higher correlation between *IQ*'s and social status scores when the choosing situations involved some degree of ability such as being a class librarian or acting the part of a doctor in a health project. In one school a girl with the highest *IQ* in the group had never received a popularity score of even as much as average in either the first or second grades, but in an election of post-master in the second grade she rose to fourth place from the top. The teacher had called attention to the need of certain qualifications for the office. In McGahan's (14) study in which coefficients in the thirties and forties were obtained all the choosing situations involved leadership functions.

It seems highly probable that the chief reason for the generally low correlations between *IQ*'s and the status scores lies in personality traits. A boy with an *IQ* of 120 in one of the schools studied is very low in status apparently because he is inconsiderate of the rights of others. He fights and pushes ahead of his turn. A girl with an *IQ* in the gifted class is in the lowest quartile in popularity apparently because she is so indifferent to other children that she has not made friends. On the other hand is a girl with an *IQ* of 101 who ranks next to the top in social acceptance in her grade. She performs well before the group, tells stories with dramatic effects, dresses well, has a cute smile, and is quite attractive. The lowest *IQ* found in any upper quartile group was 77. This was

the score of an over age boy who was very poor in his school work but was quite active and capable on the playground. In this capacity he won the admiration of other boys who gave him most of his votes.

It is quite possible (and some evidence shows it) that the relation between intelligence and social acceptance increases with age.

E. Academic Achievement and Social Status

In his study of sociability of elementary school children Hsia (10) found a correlation of +.33 (.46 after correction) between his composite sociability scores and reading achievement on a standardized test. Hardy (8) found an r of +.36 between total educational achievement and pupil choices of friends in grades three to six.

The reading data in the Denton study was obtained from the Gates *Primary Reading Test* which was given in its different forms in September, January, and May. Miss Nellie Griffiths and her student assistants in the Reading Laboratory were responsible for administering and scoring the tests. The findings in relation to social acceptance are summarized in Table 5. The same procedure

TABLE 5

Relation Between the Combined Social Status Quartiles in Three Second Grades in Denton Schools and Reading Achievement

Combined social status quartiles	Ave. R.A. Sept. 1939 71 cases	Ave. R.A. Jan. 1940 64 cases	Ave. R.A. May 1940 49 cases	Ave. gain Sept. to Jan. 63 cases	Ave. gain Jan. to May 51 cases
4	7 yr.-4m.	8 yr.-7m.	8 yr.-7m.	12 m.	7.5 m.
3	7 yr.-6m.	8 yr.-3m.	8 yr.-7m.	10 m.	6 m.
2	7 yr.-1m.	8 yr.-1m.	8 yr.-7m.	9 m.	8 m.
1	6 yr.-10m.	7 yr.-6m.	8 yr.-3m.	7 m.	7 m.

is followed as in the previous section of combining the status quartiles of all three schools.

Turning to the data in Table 5 on reading ages (R. A.) it can be seen that in September there was some differentiation between the status quartiles in reading ability, but it is not entirely consistent since the highest average is in the third quartile. However, the two upper groups have an advantage of 5.5 reading age months over the two lower groups. In January there is a very consistent variation

of average reading ages with the social status quartiles. The differences run 7 months, 2 months, and 4 months between the first and second, the second and third, and third and fourth quartiles respectively. The two upper groups have an advantage of 7.5 reading age months over the two lower groups. In May there is no variation of average reading ages with the social acceptance quartiles, except that the lowest quartile is four months below the other three. Why this should be is not known for certain. There is a cue, however, in the two succeeding columns in Table 5 on average gains. It will be noted first that the gains made from September to January are greater than the gains from January to May in the three upper groups. It will also be noted that the September to January gains vary consistently with the status quartiles. Apparently what this means is that during the first semester, as compared with the second, the pupils were more differentiated in their reading progress and that on the whole the more popular children made the greatest reading gains. During the second term the better readers were not able to maintain the rate of progress established during the first semester on easier material—so there was a levelling off of gains and consequently a poorer relation between social success and reading progress.

A number of correlations were run between the scores obtained in particular choosing situations and reading ages. These ran: +.04, +.08, +.10, +.41, +.54, and +.67. The last three coefficients, which are so noticeably higher than the others, were obtained in the first instance in January and in the other two in the late spring. A possible explanation of this rise in the size of the coefficients is that as the school year progressed there was a tendency for the better readers to become a little more socially adequate and thus to be better accepted. This may seem to be a contradiction to what has just been said in interpreting the data of Table 5, but there is no real contradiction since the quartiles in Table 5 are based on the average status scores throughout the year, whereas the three correlations of +.41, +.54, and +.67 were obtained on choices made in January, April, and May only.

In School *A* the teacher preferred to give the *New Stanford Achievement Test* rather than the Gates *Primary Reading Test* at the close of school in May. A correlation of +.46 (22 cases) was found between the composite score on this test and the average social status for the full year. Although this cannot be said to be

high, nevertheless, when the size of other coefficients that have been obtained between any one factor and social acceptance scores are considered, it must be concluded that good general academic competence in this group was quite definitely related to social status.

F. Various Physical Characteristics and Social Status

Hardy (8) found in his study of elementary school children that his upper group in popularity was rather decidedly superior to the general population studied, and especially to the unpopular group, in both athletic skills and general health. The most popular ones had fewer defects of teeth, tonsils, adenoids, and posture. Height and weight had no significant relation to social acceptance. Physical attractiveness was much more characteristic of the popular than of the unpopular group.

Hicks and Hayes (9), studying adolescents, found practically no difference between their four groups on the basis of social competence and physical measurements. In eye defects the upper group was in a worse condition than the lowest group, the respective figures being 17 and 7 per cent.

In the Denton study data were obtained on the following physical factors: teeth cavities, enlarged or inflamed tonsils, certain eye and ear defects, dark adaptation, basal metabolism, and height and weight.

The data on teeth cavities were obtained from examinations made by two practicing dentists in the city of Denton who came to the schools in November, 1939, for this purpose. Likewise, the data on

TABLE 6

Relation Between the Combined Social Status Quartiles in Three Second Grades of Denton Schools on the One Hand and Teeth and Tonsil Defects on the Other

	Teeth Cavities (80 cases)		Tonsil defects (76 cases)		
Combined quartiles for three schools	No. of cavities	No. of children without cavities	No. enlarged	No. inflamed	No. of children without tonsil defects
4	69	5	11	10	9
3	50	6	7	3	11
2	49	5	9	7	9
1	69	1	10	6	8

tonsil defects were obtained from examinations made by a practicing eye-ear-nose and throat specialist in Denton. The results of these surveys in relation to the combined status quartiles are given in Table 6.

A little inspection of Table 6 is all that is necessary to see that there is not a clear differentiation between the four groups on the basis of the number of teeth and tonsil defects. The number of cavities is exactly the same in the lower and upper quartiles and the number of tonsil defects is a little more in the fourth (highest) than in the first quartile. It is probable that more relation might be found between teeth and tonsil defects and social acceptance if the doctors had differentiated more between serious and mild cases of defects. This will be done another year.

The data in respect to eye defects were obtained by use of the Bett's Telebinocular.[3] The purpose of this instrument is to screen out the cases which have deficiencies in visual power or in certain types of visual functioning such as efficiency, muscular imbalance, and fusion. The use of this instrument was under the direction of Miss Nellie Griffiths of the Teachers College faculty. In summarizing the results of these tests each child was rated as "satisfactory," "questionable," or "unsatisfactory." The latter rating was given only when there was no doubt whatever of the child's visual defect as measured by the telebinocular. The results of the visual tests in relation to the combined social status quartiles of the three Denton schools is given in Table 7.

TABLE 7

RELATION BETWEEN THE COMBINED SOCIAL STATUS QUARTILES IN THREE SECOND GRADES IN DENTON SCHOOLS AND VISUAL DEFECTS

Combined status quartiles	Visual defects (69 cases) Number satisfactory	Number questionable	Number unsatisfactory
4	7	10	0
3	6	7	4
2	6	6	6
1	7	3	7

Inspection of the data on visual defects in Table 7 shows that the four status groups are approximately equal in number rated

[3]Distributed by Keystone View Co., Meadville, Pennsylvania.

satisfactory. The highest group has the worst record on the rating of questionable, but since the actual condition of the eyes of these pupils is uncertain, it would not be wise to conclude anything from these figures. It will be observed that none of the most popular children had a rating of definitely unsatisfactory, and that the figures for the other three quartiles vary concomitantly with degree of popularity. It would seem that this finding has some significance, indicating as it certainly does within the limits of the cases involved, that being free from clearly unsatisfactory eyes is an asset in attaining those characteristics which make for social success. This is exactly contrary to the report on this point given by Hicks and Hayes (9) previously cited. However, the data obtained by these workers are not comparable with that obtained in Denton because of differences in method. In the Hicks and Hayes study social success was determined by teacher ratings and, more important, the factor of scholarship was included in the composite scores. It is reasonable to suppose that eye defects would have more relation to such leadership or social success scores than to popularity scores obtained by pupil choices alone. Furthermore, the age difference between second grade and junior high school pupils is no doubt important.

Under the supervision of Miss Nellie Griffiths 80 children were tested with the Audiotest.[4] The purpose of this machine is to locate those children who seem to have enough hearing loss to interfere with comfortable learning in the classroom. However, it turned out that only four children had unsatisfactory hearing. Since there was so little difference between the children in this respect it is not possible to show any relationship between the results with the Audiotest and the status quartiles.

Data relative to the dark adaptation of the Denton children were obtained by use of the biophotometer manufactured by the Frober-Frabor Co. The use of this instrument was under the supervision of Dr. Florence Scoular, Director of the Department of Home Economics of the North Texas State Teachers College. All the measurements were made during the months of December, January, and February. The purpose of the biophotometer is to measure the rapidity at which visual purple is regenerated. Vitamin *A* is necessary to this regeneration. Thus it is a measure of the Vitamin *A*

---

[4]Made and distributed by Educational Laboratories, Inc., Brownwood, Texas.

status of a subject. The technique used in the tests was that described by P. C. Jeans, E. Blanchard, and Zelma Zentmire and the norms used were those established by these workers.[5]

The basal metabolism scores were obtained only on the children in School *A*. These measurements were made by Miss Hattie Taylor, the technician in Teachers College Hospital. The machine used was the Sanborn Metabolism Tester. Nearly all the tests were made during the months of February, March, and April. In order to acquaint the children with the apparatus before taking the test, each child was brought to the hospital a day or so before his turn and allowed to try on the nose and mouth pieces, and was told what he would be expected to do. The children were brought to the hospital around eight o'clock in the morning without having eaten. All other standard conditions were followed. In case a child did not perform well on the test he was required to repeat it at a later date. There were only three such cases.

A summary of the data on dark adaptation and basal metabolism in relation to the social acceptance quartiles is given in Table 8.

TABLE 8

RELATION BETWEEN THE COMBINED SOCIAL STATUS QUARTILES OF THREE SECOND GRADES IN DENTON SCHOOLS ON THE ONE HAND AND DARK ADAPTATION AND BASAL METABOLISM (SCHOOL *A* ONLY) ON THE OTHER

Combined social status quartiles	Dark adaptation (77 cases) No. at norm	No. below norm	No. in norm range	Basal metabolism School *A* Only—(23 cases) No. above +.10	No. below —.10	Average
4	2	18	4	2	0	+5.3
3	3	16	2	2	1	+5.4
2	3	16	3	3	0	+11.1
1	1	18	5	1	0	+5.3

Only a brief examination of Table 8 is necessary to see that there is not a consistent relation between either the dark adaptation scores or the basal metabolism scores and the social status quartiles. All this can mean is that these particular physical factors cannot be said to have played a significant rôle in determining the social acceptance of these children.

Correlation coefficients were run between the basal metabolism

[5]P. C. Jeans, Evelyn Blanchard, and Zelma Zentmire. "Dark adaptation and vitamin *A*." *J. Amer. Med. Asso.*, 1937, **108**, 451-458.

ratings and the total status scores for both the first and second grades. In the first grade (27 cases) this proved to be —.01. In the second grade (25 cases) the *r* was +.019. Both these figures bear out the data of Table 8 in showing that there is practically no relationship between the basal metabolism scores and the attainment of social status.

The correlation between 19 carry-over cases from the first to the second grade on the metabolism test proved to be +.18. In the first grade the scores ranged from —12 to +21 with a mean of +5.52. In the second grade the range was from —11 to +18 with a mean of +7.08.

Data on height and weight were available in Schools *B* and *C*. In both schools the measurements were made in February, 1940. Correlations were run between both these measures and the total average status scores. In respect to height the *r*'s were —.09 and +.08 respectively in the two schools, and in regard to weight the *r*'s were —.05 and +.41 respectively.

G. Personality Traits and Social Status

Numerous studies have reported data on the personal traits related to social acceptance. Typical of these are those of Murphy and Murphy (13), Hicks and Hayes (9), Flemming (6), Burks (2), Thomas and Young (20), Pechstein and Mann (17), Tryon (21), and Dimock (4). From these studies a list of such traits as the following are usually found to be of greatest importance: coöperation, courtesy, cheerfulness, modesty, intelligence, truthfulness, control of temper, unforced humor, adaptability, and sympathetic interest in others.

The data on the personal traits of the Denton children were obtained from teacher judgments. At the end of the school year the teachers were asked to give a brief description of the outstanding traits of each child. These were given orally and taken down by the writer. In some respects a numerical rating scale would have been better, but this written-portrait method had the advantage of bringing out the most emphatic characteristics of each child. Therefore, the data presented under this heading may be considered to represent rather marked degrees of the traits listed. At the time the teachers gave their descriptions they did not know the final status scores which the children had, but they naturally knew a good deal about who the most and least popular pupils were. In Table 9

TABLE 9

RELATION BETWEEN PERSONALITY TRAITS AS OBSERVED BY TEACHERS AND THE COMBINED SOCIAL STATUS QUARTILES OF THREE SECOND GRADES IN DENTON SCHOOLS

Traits	No. of cases in status quartiles 1	2	3	4	Traits	No. of cases in status quartiles 1	2	3	4
Cheerful disposition	0	2	5	10	Passive—no contribution	6	1	1	0
Depressed disposition	5	3	1	0	Poor sport—blames others	4	2	2	0
Aggressive in social situations	3	1	5	8	Serious-minded deliberate	0	5	3	0
Generous considerate sympathetic	1	1	3	8	Good sense of humor	0	0	3	5
Able and confident before group	0	2	8	13	Shows off—strives for attention	2	4	4	2
Timid and incompetent before group	16	12	5	5	Boy effeminate—tied to mother	1	4	3	2
Good appear-ance—attractive	1	4	7	15	Babyish—immature	2	2	4	1
Poor appear-ance—un-attractive	6	3	2	0	Fights—quar-rels—overrides others	5	3	2	1
Dominating—bossy	2	1	2	4	Fussy-Whiney—complains, pouts	3	3	2	1
Cooperative in group—good sense of duty	0	3	1	15	Dishonest in tests—steals	1	1	0	1

is given the list of behavior characteristics mentioned by the teachers together with a summary of the number of children in each quartile (when all three schools were combined) who were designated as having a particular trait.

An examination of Table 9 shows that the social status quartiles are most clearly differentiated in respect to the following traits: Having a cheerful rather than a depressed disposition, being aggressive in social situations, being generous and considerate of others, being able and confident before the group, having an attractive personal appearance, being coöperative in group endeavors and having a good sense of duty, having a contribution to make to the group, having a good sense of humor, an avoidance of fighting, quarreling, and over-bearing conduct. These results are in line with the

findings of other studies. It can be noted that the trait which most clearly differentiates the highest from the lowest group is the one designated "coöperative in group—good sense of duty." This is of particular interest because very much the same trait, referred to as "adaptability in coöperative activities in work and play" was found by Hardy (8, p. 375) to show the largest variation between his most popular and least popular groups. Apparently this is a trait of paramount importance in attaining social acceptance.

A number of special points in Table 9 are worth mentioning. It will be observed, for instance, that a few of the children in the highest quartile in popularity are designated as having some personality traits which are generally considered to be very unfavorable such as: being dominating or bossy, being a show-off and striving for attention, being effeminate in case of a boy, being babyish, fighting, pouting and dishonesty. There is really nothing surprising about this since throughout life we have to choose our friends from among imperfect individuals on a relative basis. So it is with children. They are liked or disliked not on the basis of one or two or a half dozen traits but on the basis of the impression they make as total individuals. A child may be correctly described as dominating and yet be well liked because of other very favorable traits. Also some children (as well as adults) no doubt like to be bossed around a little by someone they admire. In certain extreme cases the domination contributes to a personality need.

Showing off and striving for attention very often defeats a child in his effort to be accepted. There were some good examples of this among the Denton children. However, in the cases of children who had this trait and were nevertheless popular it was observed that when they got attention they usually had something to contribute which the group admired. Furthermore, there are different ways of striving for attention—some of which are very antagonizing while others are only mildly so, if at all. Again, it is a case of children reacting, not to an isolated trait, but to a total individual who is demanding attention. Finally, on this point it must be remembered that showing off is not always regarded in the same light by children as by teachers.

The effeminate boys who fall in the upper quartile of status got nearly all their choices from girls. They preferred the company of girls and played with them at play periods.

Traits which the teachers called babyish or immature apparently

seem cute or "sweet" to some children. Their submissiveness and weakness no doubt appeal to those children who want to make sure that they can feel dominant, or at least competent, in their inter-personal relationships. It is probable that children are not nearly as sensitive to traits of immaturity as are adults who are trained to look for such things.

The child who got in the upper quartile in spite of some tendencies to fight and quarrel is a boy who received most of his votes from other boys apparently because of his aggressive competence on the playground.

The child in the highest popularity group in the category "fussy-whiney-complains-pouts" is a girl who pouts rather often when she is thwarted, but who is also very nice and pleasant on other occasions, is very bright, dresses exceptionally well, and is admired for her all-round competence in the group.

The one case of a child who is very popular, even though she is designated by the teacher as dishonest, is a girl who has an *IQ* of 100 and is not as able in her school work as the other girls who are her friends and who are brighter than she is. The teacher caught her cheating in written work a few times. She was next to the top in popularity in her group.

These cases of very popular children with unfavorable personal traits, as well as other data in Table 9, show the fallacy of such easy generalizations as: "Everybody likes a cheerful person," "You can't be popular if you are timid,". "Nobody likes a bossy person," or "Nobody likes a person who pouts." It is one thing to say that everybody likes cheerfulness in a person, but this is not equivalent to saying that everybody likes a cheerful person. Likewise it is probably true that nobody likes the trait of pouting, but it does not necessarily follow that the person who has this trait is disliked. Each individual is a unique whole and is judged by the total impression he makes. He is not judged on a part by part or trait by trait basis. Gestalt psychology in particular has emphasized that any functional whole is more than a mere sum of its parts. The meaning of any trait in a child cannot be determined simply from knowing that he has that trait. It takes on meaning only when it is known how that trait functions in the total matrix of his personality. This is why rating scales are so unrevealing unless accompanied by much supplementary data.

H. Teacher Ratings and Social Status

Near the end of the first semester in the second grade the three teachers in the Denton study were asked to place their pupils in five groups as follows: *A*—most sociable, *B*—above average in sociability, *C*—average in sociability, *D*—below average in sociability, and *E*—very unsociable. Sociability was to be judged according to the following criteria:

1. Ability to work harmoniously with fellow pupils.
2. A "good mixer."
3. Number of social activities participated in.
4. Degree of tact in getting along with children.
5. Number of school activities participated in.
6. Any other activities which make a boy or girl popular both in school and out of school.[6]

The teachers were told that they need not put the same number in each group, and, as a matter of fact, they did not. The tendency was to place most of the children in Groups *B* and *C*. In order to determine the extent of agreement between the teacher ratings and the pupil choices it was decided that it would be best to use only the results of the choosing situation closest in time to the teachers' ratings. This was either a December or a January choosing. The scores in these situations in each school were divided into five equal groups from highest to lowest. The data in Table 10 show the extent of agreement between these groups and the teacher ratings.

TABLE 10

Relation between the Social Success of Second Grade Children As Determined by Teacher Ratings and Their Social Success by Pupil Choices

Social success group by teacher ratings	Per cent of group in agreement with pupil choices			Average amount of variation of teacher rating groups from pupil choice groups		
	School *A*	School *B*	School *C*	School *A*	School *B*	School *C*
A	.75	.50	.50	.50	1.60	1.16
B	.20	.42	.33	1.40	.57	.89
C	.16	.25	.11	1.33	1.37	.89
D	.32	.66	0	.83	.33	1.
E	0	.66	.33	1.66	.33	1.

[6]This is a slight modification of the criteria set up by Jui-Ching Hsia in his *Study of the sociability of elementary school children.*

In examining Table 10 it can be seen that the most consistent degree of agreement between the teacher ratings and the pupil choices is in Group *A*—the most socially successful children. These figures were arrived at by finding the per cent of the names placed in each group by the teacher which corresponded exactly with the names in corresponding groups by pupil choices. Below Group *A* the degree of agreement is not high and it is quite variable from group to group. However, the data in the rest of the table show that the extent of the disagreement between teachers and pupils was, on the whole, not great. When the average deviation between the two is taken for each group, it can be seen that the greatest amount of disagreement is only a little more than one and one half groups. More than half (9 out of 15) of the average deviations are only to the extent of one group or less. In all three schools only three children who were placed in the two upper groups by the teachers fell in one of the two lowest groups on the basis of pupil choices. In only one case did a child placed in Group *A* by a teacher fall in Group *E* by pupil choices. This data would seem to justify the conclusion that, on the whole, the teachers and children did not differ greatly in their selections of socially successful pupils. Certainly nothing like perfect agreement would be expected, since, in the first place, teachers are not adequately trained to notice personality patterns, and, in the second place, we know that children frequently choose as friends other children who have serious personality weaknesses from an adult point of view.

At the end of the school year in May the teachers were asked to make an estimate for each child in respect to whether they considered him to have gained or lost or remained approximately constant in social status during the time she had the class. In School *A* this was the entire school year, but in Schools *B* and *C* it was only the second semester. The estimates were obtained by having the teachers check each child on a scale which ran from —10 per cent to +10 per cent. The instructions were to check each child according to the extent he was considered to have gained or lost, if any, on the basis of the voting in the choosing situations. At this time the teachers did not know the actual scores of the pupils. Correlation coefficients were run between the per cent scores awarded by the teachers and the actual per cent scores determined by subtracting each child's score in May from his average for the full year or full semester. For School *A* this coefficient was +.19, for School *B* it was —.11, and for School

C it was —.10. The higher degree of relationship found in School *A* was no doubt due to the teacher's having had the pupils for the full year rather than for just one semester. It is clear, however, that all three coefficients are very low—the last two being negative. These low relationships may be regarded in two different ways. In the first place, if the pupil choices are taken as the criterion it would be concluded that these three teachers made poor judgments about the degree of progress their pupils made in social acceptance. From this standpoint the teachers over-estimated the amount of gain which the children actually made. Analysis of the data shows that the most typical difference between teacher ratings and the pupil choices was that the teachers awarded the pupils higher gains than they had actually made. This was also true of the teachers' placement of their pupils in the various groups previously discussed. No doubt it is a very common tendency for teachers, as well as others engaged in efforts to help people, to over-estimate the amount of good that has been accomplished. However, there is a second way of looking at these differences. It may very well be that the teacher estimates of social progress are more valid than the pupil choices. It is quite probable that a conscientious teacher is more sensitive to a child's improvement in his social relationships than are other children in the group. She is watching for it more and has a more adequate standard from which to judge. In order to be most certain of a child's progress in his social relationship it would be necessary to have a rather high degree of agreement between teacher estimates and pupil choices.

I. Sex Differences in Social Status

A number of studies have reported data on sex differences in relation to social competence. In Hsia's (10) study it was found that girls were slightly favored over boys in composite sociability score. Hardy (8) found in his highly selected group of most popular children in the elementary grades that 68 per cent were girls, whereas only 30 per cent of the least popular children were girls. Hicks and Hayes (9) found in studying junior high school pupils that teachers consistently rated girls higher than boys in health habits, study habits, and relations to other pupils. Witty (22) found that senior high school teachers rated "only" girls higher in nine character traits than they did "only" boys or both boys and girls in intermediate family positions. Smith (18) found in studying

the opinions of boys and girls of each other between 8 and 15 years of age that with increase in age the boys had a progressively poorer opinion of the girls, and the girls a progressively better opinion of boys. Boys, on the other hand, had a progressively better opinion of themselves, and girls a progressively poorer opinion of themselves.

The data from the Denton schools on sex differences is in line with most of that referred to above, in that the girls received consistently higher scores than the boys. This point is shown in Table 11.

TABLE 11

SEX DIFFERENCES IN TOTAL AVERAGE PER CENT SCORES IN SOCIAL STATUS IN THREE SECOND GRADES OF DENTON SCHOOLS

School	No. of boys	No. of girls	Ave. for boys	Ave. for girls	Difference in favor of girls
A	9	14	.035	.05	.015
B	25	13	.025	.036	.011
C	15	16	.026	.042	.016

It can easily be observed from the above table that the girls in each school received a total average per cent score which was more than one per cent higher than that of the boys. It is significant that this holds true regardless of the proportion between the sexes. In School *A* the girls outnumbered the boys, in School *B* the boys outnumbered the girls almost two to one, and in School *C* the numbers were approximately equal. The results in School *B* are particularly interesting because there is a strong tendency for children of this age level to give their votes largely to members of their own sex. On this basis the boys had a much better chance to receive votes than did the girls, and yet the girls got the higher average, but not quite to the extent that was true of girls over boys in the other two schools. In all three schools the top ranking child in status was a girl and in two of them the top three were girls. Out of the 23 pupils who composed the highest quartile of all three schools combined only eight were boys.

These facts, together with the body of evidence previously referred to from other studies, certainly seem to point to the conclusion that girls are better socialized than boys. Whether this is due to a real sex difference, to the more rapid maturity of girls, to certain kinds of teaching, to the school being more adapted to girls than to boys, to a greater number of women teachers, or to all these factors combined is not known. Certainly if these findings are sub-

stantiated by more extensive research they constitute a challenge to educators to try to construct a school program which does a better job of promoting the social adequacy of boys in particular. Not the least of the considerations involved is the fact that some of the most analytical studies of delinquency have shown that such factors as a sense of inferiority and social incompetence are very frequently the psychological origins of anti-social conduct. This finding, together with the results reported above, may have some relation to the fact that delinquency is about six times as common in boys as among girls.

There may, however, be another side to this matter as indicated in the study cited by Smith (18) in which both boys and girls between ages 8 and 15 had a better opinion of boys' personality traits than they did of these traits in girls. Does this mean that the kinds of social adjustments made by boys are really superior to those made by girls? Or does it mean that the boys have attained a higher degree of integrity even though they are not so sociable. Or are the findings in the Smith study due to the young people gradually absorbing as they grow older the rather general tendency in our culture to play up men over women and masculine traits over feminine traits? Or is the evidence on this point too scanty as yet to be taken seriously? It seems safe to say, however, that educators as a whole have slidden too easily over considerations of psychological sex differences.

As previously stated there is a very strong tendency for boys and girls of this age level to confine their choices to their own sex. The data in this respect for the Denton study may be summarized by the following statements: In School *A* the average number of boys chosen by boys throughout the school year as compared with girls was in the ratio of 8 to 5 and the average number of girls chosen by girls as compared with boys was in the ratio of 13 to 7. In School *B* the average number of boys chosen by boys as compared with girls was in the ratio of 15 to 5 and the average number of girls chosen by girls as compared with boys was in the ratio of 12 to 4. In School *C* the average number of boys chosen by boys as compared with girls was in the ratio of 10 to 2 and the average number of girls chosen by girls as compared with boys was in the ratio of 14 to 7. These figures show how strong is the tendency for both boys and girls to choose within their own sex.

J. Certain Home Factors and Social Status

Most investigations have shown that the socio-economic status of the home is one of the factors most closely associated with social success in school. Hardy (8) found an r of $+.46$ between social recognition and home and neighborhood ratings. Hsia (10) obtained an r of $+.30$. Fenalson and Hertz (7) found that college students whose fathers' occupations fell in the upper brackets on the Minnesota occupational intelligence scale had considerably fewer feelings of inferiority than did those whose fathers' occupations fell in the lower brackets.

Data on the home backgrounds of the Denton children are not yet complete but there are some things to report for Schools B and C. Mrs. Flora Patrick (16) made a study of the relation between the social acceptance scores of the second grade children in School B and the occupational status and sociability of the parents of these children. Using the Barr scale to determine occupational intelligence status she found a correlation of $+.52$ with the social status scores of 33 children for the full term. The sociability of the parents was determined by having them answer a questionnaire (through a personal visitation by Mrs. Patrick) bearing on the extent of their participation in various community groups such as lodge, church, P.T.A., committees, clubs, and sports. Participation during the past two years was counted. Extra credit was given for holding an office. Altogether there were 50 items in the questionnaire. Between the total sociability scores on this questionnaire and the social acceptance scores (33 cases) an r of $+.62$ was obtained. This study is not sufficiently extensive to be very significant but it does offer a promising lead for further research. It would certainly be important information for parents if it could be definitely shown that their own social adequacy and interest is one of the most important factors determining the social success of their children in school.

In School C the vice-principal, Mr. J. D. Parnell (15), made a survey of homes of the second grade children using the *Minnesota Home Status Index* by Alice M. Leahy. A correlation was run for 30 cases between the cultural status index found from this scale and the social status scores for the full year. The r was found to be $+.63$. This is in line with the results of other studies which have emphasized the importance of a child's home background in determining his social acceptance at school.

K. Size of Family and Social Status

Hardy (8) found an r of —.30 between social recognition and family size. This shows a tendency for the best liked pupils to come from smaller family units. Fifteen per cent of the most popular children were "only" children. The Denton data are in line with Hardy's findings. The relation of the family size of 81 second grade pupils to their social status is given in Table 12.

TABLE 12

Relation between the Combined Social Status Quartiles in Three Second Grades in Denton Schools and Size of Family

Combined social status quartiles	No. of brothers	No. of sisters	Total number siblings	No. of siblings within 4 years of age of sec. grade child	No. of "only" children
4	7	21	28	8	8
3	18	18	36	12	3
2	20	19	39	19	4
1	20	30	50	23	3

It can be seen from the above table that there is a consistent tendency for the more popular pupils to come from smaller family units, particularly in respect to total number of siblings and the number within four years of the age of the second grade child. The lowest group in social acceptance had 13 more brothers and 9 more sisters, or a total of 22 more siblings with whom to associate than did those in the highest popularity group. The figures on the number of brothers and sisters within four years of the age of the second grade child were included because it is not likely that a seven year old child would have much in common with a brother or sister who is more than four years older or younger than himself. The data on this point show a very consistent tendency for the more socially successful children to have fewer brothers and sisters near their own age than the less popular children. The four groups are more consistently differentiated in this respect than any other aspect of family size. The chief point of interest in the number of "only" children is that there are at least twice as many of these in the highest quartile in status as in any of the other three groups. Five of the six most popular pupils in School A were "only" children. In

each of the three schools the highest total average per cent score in social acceptance was held by an "only" child.

From knowing the children in the three second grades involved in this study it seems safe to say that the chief reason why the child from the smaller family units is, on the whole, more successful in achieving social acceptance is that he has superior home advantages.

The idea that the "only" child is especially handicapped in his social adjustments probably arose in an earlier day when transportation was much more difficult and infrequent than it is today. There is no reason why an "only" child in any typical American community at the present time cannot have as many playmates and get as much socialization as any other child. Furthermore, there are certain unfavorable factors at work in many large families. If the income is low there may be a rather constant struggle between the children to secure clothes and other advantages. This can easily result in selfishness and other unfortunate traits. Sometimes older children in a family assume practically all the responsibilities leaving a younger one without normal duties, or he may be waited upon and babied by the older ones. It is probably safe to say that the degree of intelligence shown by parents in handling their children is a much more important factor than the number of children in the family.

L. Chronological Age and Social Status

In previously published studies chronological age has not been found to have much relation to popularity within particular groups. Data on this point for the Denton study were gathered first by combining the social status quartiles for the three second grades and then finding the average age of each group. When this was done it was found that the average age (Jan., 1940) for the highest quartile was 93 months, for the third quartile it was 94 months, for the second it was also 94 months, and for the lowest group the average age was 95 months. On the whole these figures do not show much relation between social status and age, but it is of interest to note that the lowest group has an average age which is two months higher than that of the top group. This shows a tendency for the over-age retarded pupils to make a relatively poor social adjustment. This point has often been emphasized in educational discussions on the wisdom of retardation.

M. Summary

Social status in this study was determined by pupil choices in five to eight choosing situations throughout the school year in the second grade.

Status was found to be quite highly concentrated in a few pupils at the top. The most popular children received their votes from a wide range of other pupils in their groups rather than from just a few close friends. There was a marked tendency for more children to choose those in the upper quartile in status than was chosen by these children in return. This shows that those who were low in acceptance were reaching out for more social contacts whereas those at the top were apparently well satisfied with their present position in the group, and so did not feel the same degree of motivation to try to bring others within their social circle.

Stability of status was quite high as shown by correlations between successive choosing situations particularly in Schools *A* and *C*. The lower correlations in School *B* were apparently due to two factors, first, the greater homogeneity of pupils in this school in *IQ*, social competence, and home background, and second, to a greater pupil turn-over throughout the school year.

The relation between *IQ* and social status was on the whole not marked, but the first and fourth quartiles in status (when the three grades were combined) showed a significant difference in *IQ* averages—being 16 and 17 points respectively in the first and second grades. The chief explanation for the lack of high correlations between *IQ* and social status seems to lie in personality traits.

Some relationship was found between reading ability and social status particularly in respect to the reading ages obtained in January and the reading gains made between September and January. It seems probable that there was a greater relationship between achievement in reading and social acceptance during the period when reading progress was the greatest. Correlation coefficients obtained between reading ages and scores made in individual choosing situations were higher during the latter part of the year, which would seem to mean that once the better readers became established (although they did not maintain their rate of gain) there was a tendency for them to rise in social acceptance. In School *A* an r of .46 was found between scores on *New Stanford Achievement Test* and status scores for the full year.

No relation was found between number of teeth cavities and inflamed or enlarged tonsils and social acceptance. No difference was found between the combined social status quartiles in number having definitely satisfactory eye sight, but there was a consistent variation between them in the number having definitely unsatisfactory eye sight—ranging from 7 in the lowest to 0 in the highest quartile. No relationship was found between social acceptance and either dark adaptation or basal metabolism. The correlation between height and weight and social status were of a zero order with the exception of the r for weight in one school which was $+.46$.

Most of the generally recognized desirable personality traits were found by teachers to be much more common among the popular than among the unpopular children. However, a number of very undesirable personal traits were listed by teachers as being characteristic of popular children, and, at the same time, some very desirable traits were accredited to unpopular pupils. This shows that social acceptance is not due to certain traits but to the total impression the individual makes on others.

On the whole the teacher judgments of socially successful pupils agreed quite well with the pupil choices. The highest degree of agreement was in respect to the most outstanding pupils. When the teachers were asked to estimate the amount of gain or loss which their pupils had made during the time they taught them, it was found that the correlations between teacher judgments and the actual gains and losses were very low.

The data on sex differences show a rather strong and highly consistent tendency for girls to receive higher status scores than boys. From this it may be concluded that the girls in these three second grades are more highly socialized than the boys. The data also show a marked tendency for both boys and girls to choose friends and companions within their own sex.

A correlation of $+.52$ was found in School *B* between the occupational intelligence status of parents as measured by the Barr Scale and the social status scores of their children in school. An r of $+.62$ was found in this school between sociability of parents as measured by a questionnaire and the social acceptance scores. In School *C* a correlation of $+.63$ was obtained between the cultural status of the homes on the *Minnesota Home Status Index* and the social status scores of the children in school. These correlations are the highest obtained between any one factor and the social success scores.

There was a rather high tendency for the more popular children to come from the smaller family units. More than twice as many "only" children were found in the highest group in social acceptance than in any of the three lower groups.

Not much relation was found between the social status scores and chronological age. It is no doubt of some significance, however, that the average age of the lowest status group was two months higher than that of the most popular group. This is probably due to the poorer social adjustments made by the over-age retarded pupils.

References

1. Brown, M. Leadership among high school pupils. *Contrib. Educ., Bur. Pub., Teach. Coll., Columbia Univ.*, 1933, No. 559.
2. Burks, F. W. Some factors related to social success in college. *J. Soc. Psychol.*, 1938, **9**, 125-139.
3. Criswell, J. Social structure revealed in a sociometric retest. *Sociometry*, 1939, **2**, 69-73.
4. Dimock, H. S. Rediscovering the Adolescent. New York: Association Press, 1937.
5. Douglas, A. T., & Johnston, F. S. Environmental factors and their relation to social adjustment. *Men. Hyg.*, 1939, **23**, 379-413.
6. Flemming, E. G. A factor analysis of the personality of high school leaders. *J. Appl. Psychol.*, 1935, **19**, 597-605.
7. Fenalson, A. F., & Hertz, H. R. The college student and feelings of inferiority. *Men. Hyg.*, 1938, **22**, 389-395.
8. Hardy, M. C. Social recognition at the elementary age. *J. Soc. Psychol.*, 1937, **8**, 365-384.
9. Hicks, J. A., & Hayes, M. Study of the characteristics of 250 junior high school children. *Child Devel.*, 1938, **9**, 219-241.
10. Hsia, Jui-Ching. A study of the sociability of elementary school children. *Contrib. Educ., Bur. Pub., Teach. Coll., Columbia Univ.*, 1928.
11. Jennings, H. H. Quantitative aspects of tele relationships in a community. *Sociometry*, 1939, **2**, 93-99.
12. Moreno, J. L. Who Shall Survive? Washington: Nerv. & Men. Dis. Pub., 1934.
13. Murphy, G., Murphy, L. B., & Newcomb, T. M. Experimental Social Psychology. New York: Harper, 1931.
14. McGahan, F. E. Factors associated with leadership ability of the elementary school child. Unpublished Master's Thesis, North Texas State Teachers College Library, 1940.
15. Parnell, J. D. The relation between home factors and social and academic progress. Unpublished Master's Thesis, North Texas State Teachers College Library, 1940.

16. PATRICK, F. L. Sociability of parents and social success of children. Unpublished Master's Thesis, North Texas State Teachers College Library, 1940.

17. PECHSTEIN, L. A. & MUNN, M. D. The measurement of social maturity in children. *Elem. Sch. J.*, 1939, **40**, 113-123.

18. SMITH, S. Age and sex differences in children's opinions concerning sex differences. *J. Genet. Psychol.*, 1939, **54**, 17-25.

19. SODERQUIST, H. O. Validity of the measurement of social traits of high school pupils by the method of rating by associates. *J. Educ. Res.*, 1937, **31**, 29-43.

20. THOMAS, W. F., & YOUNG, P. T. Liking and disliking persons. *J. Soc. Psychol.*, 1938, **9**, 169-187.

21. TRYON, C. McC. Evaluations of Adolescent personality by adolescents. *Monog. Soc. Res. Child Devel.*, 1939, **4**, No. 4.

22. WITTY, P. A. Only and intermediate children in the senior high school. *J. Exper. Educ.*, 1937, **6**, 180-186.

Department of Education
North Texas State Teachers College
Denton, Texas

The Journal of Genetic Psychology, 1942, **60**, 307-320.

PIAGET'S QUESTIONS APPLIED TO A CHILD OF KNOWN ENVIRONMENT*

Department of Psychology, Louisiana State University

WAYNE DENNIS

A. INTRODUCTION

Piaget's studies in child thought may initiate a new and fertile trend in modern child psychology. Elsewhere (1) I have called attention to the fact that Piaget has had many predecessors who have reported the same facts of human nature which he discovers and finds interesting. Nevertheless, Piaget has certainly gone further in the direction which he travels than has any previous inquirer.

Piaget's thesis is that thought passes through distinctive stages before adult ideas are attained. It is not that the child grasps adult notions with different degrees of adequacy, as the prejudice toward quantitative measurement would lead most modern investigators to suppose, but rather that the child develops ideas which differ qualitatively from those of the educated adult—ideas which the older person may or may not remember as having been possessed by himself at an earlier age.

Now the ideas of the adult—the notions of the "adult stage"—are doubtless acquired by each person through contact with members of older generations. These ideas have been developed in the course of history; they have not been developed independently by the individuals who now possess them.

The intermediate stages in the development of ideas in the individual, as described by Piaget, are truly intermediate between the adult conceptions and the child's earliest conceptions. The intermediate stages are, in all likelihood, compromise formations due to the conflict between the child's own views and those expressed by adults. If I understand Piaget correctly, the children of a society whose adults hold conceptions different from our own might go through intermediate stages different from those which he outlines for French and Swiss children.

*Received in the Editorial Office December 5, 1940.

But the primary stage in the development of any conception stands apart from the intermediate and final stages by being slightly, if at all, related to the culture into which the child is born. The primary conception—whether of names, or of dreams, or of any other phenomenon treated by Piaget—is developed by the child himself. It is influenced by environment only in the sense that it comes from the child's perception of and reflection upon the scene in which he lives. It is affected, Piaget holds, to a very slight extent by the world of adult ideas, for the child at this age is either not in contact with adult ideas or does not understand them.

These primary or autogenous ideas, as I shall call them, are believed to be derived anew each generation and, in our society at least, are discarded with each generation. They represent an intellectual product not in accordance with our intellectual traditions. The existence of such ideas indicates the existence of a human nature which is not solely dependent upon contact with an historical culture.

I have presented this discussion of Piaget's system, not because his ideas are unfamiliar, but because it needs to be emphasized that the crucial portion of his hypotheses are the primary stages of the child's conceptions.

If it be assumed that the most vital part of Piaget's writings is the notion that the child, by means of its own psychological processes, develops certain characteristic elementary conceptions, it is important that data be obtained from very young children whose home instruction is known to the investigator. Piaget has worked exclusively with school children, who do not fulfill this requirement. The requirement can seldom be met unless the investigator is living in the same household as the child.

It is for this reason that I present data obtained from a single subject, my daughter. From birth, Jan. 22, 1930, through the period here reported, *M* (the subject) spent practically all of her waking time in the presence of one or both of her parents. We made no attempt to induce an educational precocity. However, our explanations of natural phenomena, which were given in response to her questions, were simple, natural science explanations, the opposite in character to most of the answers here recorded. Questions on our own part such as are here dealt with were avoided until the time of the recorded examinations. The early age of the subject at the

time the study was conducted (most of the answers here reported were obtained before the third birthday) is further evidence for the spontaneous character of her concepts. Lastly, I would point out that most of the questions were again asked after an interval to check for consistency of answers.

With *M* all of the main topics of *The Child's Conception of the World* have been taken up repeatedly, although none too systematically, through her seventh year. For reasons given earlier, it seems necessary to present in full only the material which bears upon the primary stages. However, her intermediate stages as well as the primary stages are in agreement with Piaget's account.

The topics will be treated in the order followed by Piaget.

B. The Nature of Thought

Piaget finds that children at the earliest age at which they can understand the question, *"What do you think with?"* state that they think with the mouth. At 2;9,[1] *M* was prepared for the question by asking *"Can you think of a dog?" "Can you think of a house?"* To both of these she answered in the affirmative. She was then asked, *"What do you think with?"* The answer, without hesitation, was "my mouth." To check the answer two other questions were then asked and the former question then repeated. The questions were: *"What do you walk with?" "What do you carry your dolls with?" "What do you think with?"* The first two were answered by "feet" and "hands" and the last was answered as before. When asked *"Where do you think?"* she *pointed* to the mouth. These procedures were repeated one month later, and again two years later, with identical results.

In addition to the word "think," other subjective words were used. At various times during the month 2;9 the following answers were obtained (this list represents *all* answers given to questions of this type) (Table 1).

It will be noticed that remembering, believing, wishing, and hating, as well as talking, are with the mouth.

C. Origin and Place of Names

Piaget finds that, in the first stage, names are an intrinsic part of

[1]Two years and nine months of age. All subsequent ages are expressed in this manner.

TABLE 1

Question	Answer
October 26, 1932:	
What do you laugh with?	dimples
What do you cry with?	eyes
What do you hear with?	ears
What do you remember with?	mouth
What do you wish with?	(touched her lips)
What do you hate with?	(touched her lips)
What do you get angry with?	Here, where the tears run down (cheeks)
What do you want ice-cream with?	tummy
October 28:	
What do you taste with?	tongue
What do you smell with?	nose
What do you talk with?	mouth
What do you chew with?	teeth
What do you feel sleepy with?	(pointed to abdomen)
What do you get scared with?	(pointed to abdomen)
What do you believe with?	mouth
What do you get spanked with?	hands
October 31:	
What do you clean teeth with?	toothbrush
What do you comb hair with?	comb
What do you laugh with?	cheeks
What do you walk with?	feet
What do you get angry with?	tears
What do you get sleepy with?	eyes
What do you feel hungry with?	(abdomen)
What do you feel full with?	(abdomen)
What do you feel warm with?	under my sweater
What do you feel cold with?	on the legs
What do you want ice cream with?	(abdomen)

things; they are somehow located in things, and names are known simply by looking at the object named.

M's early answers seem entirely in accord with this statement. When questioned at length at 2;9 with regard to a lamp, she stated that the name of the lamp was on the lamp, that the lamp got its name by itself and from itself. Likewise, her own name was in or on (she pointed) her "tummy" and came from her "tummy." (This was in answer to the question, *"Where did your name come from?"*) She stated that a curtain is called a curtain because it is a curtain. Its name is on its "tummy." Likewise, with the name of a jacket. *"The jacket cannot be called trousers,"* she replied to a suggestion, *"because it is not trousers."* These answers are very similar to those cited by Piaget.

An incident occurred at 2;10 which showed the reappearance of her earliest notion of names after an experience during which she observed the real source of names. At the laboratory the arrival of a pair of monkeys in which she was much interested gave rise to a long discussion as to what names should be given them. A few hours after the names were given, I asked, *"Where do names really come from?"* She answered, "You give them names, you are the name man." *Can Mamma give names?"* "Yes." We explained that her mother had given her name to her. Nevertheless, ten days later the following record was obtained: *"Where did your name come from?"* "My tummy." *"Who gave you your name?"* "Mamma." *"Where did Mamma get it?"* "She got it from my tummy and gave it back to me."

Further questioning showed that she still held that the names of all objects which we mentioned (broom, dolls, lamp) were on the respective objects. We asked her a hypothetical question: *"If we gave you a new doll, what would you call it?"* Without hesitation she replied, "Patsy." *"Where would the name come from?"* "From Patsy's tummy."

D. Dreams

According to Piaget the conception of dreams at the time when they are first distinguished from other experiences is that they are objective. They come to the dreamer and take place in the room before him. They are not conceived of as a product of the dreamer.

No questions concerning the nature of dreams were asked until 4;7. When asked what she dreamed about, *M* stated she dreamed about everything she saw in the picture books and in the movies. She mentioned animals in particular. When asked where the animals were, she apparently understood that I asked where they went at the close of the dream, for she said, "Home." Asked *"Where do you see them?"* she answered, "Here and there," and pointed about the room. She was asked whether her eyes were open or shut when she saw them and she replied, "Sometimes open and sometimes shut —I see the funny ones best when my eyes are open." I asked, *"Could the animals come in if the doors were closed?"* The answer was "Yes, just like fairy animals."

Some further questions were asked at 4;11, and the answers confirmed the objectivity previously reported. The dreams were said

to come from the movies and from store windows and to be in her eyes or in her bed.

E. Consciousness

Piaget reports that the young child conceives of all things which are in any way active or useful as conscious. The child, of course, is not asked to use the word "conscious" but rather the words "know" and "feel."

A résumé of questions and answers on this topic during the month 2;9 is given in Table 2.

TABLE 2

Question	Answer
Could the clouds feel if you pinched them?	No[2]
Can the clouds feel the wind?	Yes
Can they feel heat?	Yes
Can the chair feel anything?	It feels my hands
Does the wall feel anything?	No
Does the moon know it moves?	Yes
Does the lamp know it is there?	Yes
Does the auto know when it runs?	Yes
Does the auto know when it is standing still?	Yes
What does the auto know with?	With its mouth—right behind the engine
Can the sun see us?	Yes
Can water feel anything?	It feels my hands. I can bring some in here in my hands, then it would feel the table.

These questions were not asked without interruption, so the apparent perseveration of "Yes" answers is not real. Many other answers could be cited to show the widespread attribution of consciousness to things at this early age.

F. Magical Practices

I do not find that Piaget has given a concise statement of what is meant by magic in general, but he has separately defined four categories of magic (2, pp. 133-134). In general, one may say that the term "magic" in the present instance refers to a control over the universe which is psychological under circumstances in which the adult considers that only a physical control is appropriate.

[2]This answer probably means that clouds are like fog, cannot be grasped, and hence cannot be pinched.

My notes on magic are relatively inadequate. At ages 2;9 and 2;10, *M* had a belief in the general efficacy of an up and down movement on her part and also of a whirling movement. In a conversation about clouds she was asked if she could make the clouds move. She replied that she could do so by going "up-down, up-down," meaning jumping up and down. Such a procedure, of course, does make the clouds appear to move and so may be classed as realism rather than as magic. At another time during the month 2;9 she volunteered, *"The sun is in the sky; I move the sun when I go up and down."* This is subject to the same ambiguity of interpretation. At 2;10, in a discussion about rain, she was asked, *"Can you make it rain?"* The answer was, "Yes, I go round and round" (illustrating). She said she could make it stop by the same method. As it was then raining, she was asked to go to an open door and perform the miracle. She went to the door, stood a few seconds, and said, *"I don't want to."*

A month later when asked if she could make it rain, she said that she could. Her method was to repeat part of a nursery rhyme about rain: *"Rain, rain, go away . . ."* It will be noted that the meaning of the rhyme is inappropriate. Apparently she understood the rhyme as a proper way in which to address the rain. At that time she said that she could not make the rain stop.

In the same month she asserted that she could make the moon move: *"I go up and around and all the way round."* This movement seemed to possess a general effectiveness by whomsoever it was practiced, for she stated that the sun made the clouds move by going back and forth and up and down, and on another occasion she claimed that the moon made the stars shine by going around and around their heads.

Whether other practices are to be classed as magical depends, naturally, upon a more precise definition of the magical, which cannot be attempted here. Nevertheless, I shall cite some of the data which would be relevant by some definitions. *M* had been putting her fingers to her ears (2;10). I asked, *"Why do you do that?"* The reply was, "So you will talk quietly." I asked again, as if not understanding, *"Why do you do that?"* and she answered, "So you won't talk so loud."

She seemed to believe firmly that either she or I could make the auto run merely by blowing the horn. Answers to that effect were

given repeatedly when she was asked, *"Can you make the car run?"* One gets gasoline, she said, only to make the car run fast.

G. The Concept of Life

If the question is put, *"Is this thing alive?"* or, *"Is this living?"* according to Piaget an affirmative answer will be given by the young child with reference to anything which is useful or in any way active.

At 2;9, *M* gave the material which is presented in Table 3.

TABLE 3

Question	Answer
Are the leaves on the ground dead?	Yes
Are the leaves on the trees dead?	No, they're good, they are not dead because they are good.
When the milk bottle is empty is it dead?	No,' the milkman takes it, (in other words, it is still useful)
Is your little doll (broken) living?	No, it's dead
Is Andy (a doll in good condition) living?	Yes

She volunteered that a whistle which she possessed was living. I asked her how she would kill it and she responded that she would kill the whistle by stepping on it. She classified as living the logs in a fireplace (not burning), house steps, unlighted matches, a watch, an auto, an unlighted cigarette, and a burning cigarette. The match was dead when it had been burned; a fly was dead when hit with a swatter; tin cans were dead when they had been thrown away; a cigarette, when it was put in the ash tray.

The interviews on the subject of life have not been given in their entirety because considerable time was spent before it was discovered that the word "alive" (not used in the material presented above), which I used interchangeably with "living," was used by *M* as a synonym for *"dead."* This naturally caused considerable confusion to both of us, but when the difference in meaning was understood all answers were quite consistent. The following set of questions and answers illustrates her consistent use of this extraordinary meaning of "alive," the origin of which is quite unknown to me (Table 4).

H. Artificialism

Piaget finds that the child at first finds the origin of everything

TABLE 4

Question	Answer
Is the watch alive?	No, it's good
Is the watch dead?	No, because it's not alive
Is the auto alive?	No, nor dead
Are you alive?	No, I'm hot
What is alive?	The bunny (which had recently died)
Is she alive and dead too?	Yes

in an act of human constructiveness or in something analogous to it. The answers given below, obtained at 2;10, are in accord with this statement, except one referring to cow's milk whose origin is properly known (Table 5).

TABLE 5

Question	Answer
Where does candy come from?	From the store
How is it made?	A candy-man makes it
How is chocolate milk made?	You stir it and stir it and stir it
How are rolls made?	(Gesturing) you make a biscuit and then it's bread
How was the sun made?	A sun-man made it
How is pie made?	You roll it and then it's pie
How are moving pictures made?	You put a curtain over it and that makes the pictures
How is milk made?	The cow makes it
How was the moon made?	A moon-man made it
How was our dog made?	Hamblin (who gave the dog to *M*) made him.

Two weeks later the same ideas were checked and almost identical answers were given. Questioned further as to who made the stars, she gave her stereotyped answer, *"The star-man; and the moon-man made the moon, and the baby-man makes babies, and the daddy-man makes daddies."* In an attempt to catch a perseverating answer, I asked again about milk and obtained the answer "cow" without hesitation.

At 4;7, inquiries concerning artificialistic creation were again undertaken. The results can scarcely be put in more abbreviated form than that of the original conversation (Table 6).

The last remark, as well as the previous one about the source of milk, shows how naturalistic answers are given in spite of artificialistic suggestion when they are known. Artificialistic answers are given when the child is required to invent an origin.

TABLE 6

Question	Answer
How was the swimming hole in the river made?	They dug a great big ditch
Who made it?	I don't know
Who made the mountain?	Some one sloshed it up
Who made the bee?	It was born from its mother
Who made the dog?	It was born from its mother
How was the sun made?	I don't know; I think bv its mother
How was the sky made?	I don't know; something did. The mountains?
How were the clouds made?	From mother clouds
How are toys made?	At factories
Who makes them?	Men
How are pebbles made?	They break off big rocks

I. Attainment of Adult Ideas

The subject of the preceding observations, most of which were made at the age of 2;9 and 2;10, was submitted to the same procedures at 4;7, 4;11, and 6;2. These sets of interrogations are not sufficient to enable me to determine the exact age at which various changes in ideas occurred, nor do they contribute much information concerning the processes by which these changes come about. The data derived from the later interrogations are valuable, however, in respect to the problem of the age at which adult ideas may be attained. Piaget, in his writings, tends, perhaps unintentionally, to give the impression that changes in concepts occur in all children at about the same age. This, of course, is contrary to the general tenor of American psychology, which is always aware of individual differences. In the particular field of children's concepts, Russell (4) using a standardized procedure (3) has shown that any stage of animism may be exhibited at any age from 6 to 13.

It may be worth while to indicate that the subject of the present report attained the adult stages of the conceptions here under discussion at a very early age. Many of the adult conceptions were developed at so early a period that if she had been questioned only upon entrance to school there would have been little or no evidence of any pre-adult answers. This fact indicates that in studies of child thought any negative results obtained from school children must be interpreted with caution. If a child possesses adult con-

cepts at six years of age, this is no indication that earlier concepts have never been developed. In *M*'s case, many early ideas had disappeared by the age of 6;2.

At 4;11 *M* still replied that she thought with her mouth. By 6;2, however, she held that thinking was done by the brain.

At age 4;11 she no longer believed that names were imprinted upon their object, but thought instead that names had been put into various books by persons who knew the correct names. As to how the correct names were first ascertained, she was uncertain. At the date of the last examination she was familiar with the fact that things have different names in different languages, and thought that the name of an object could be changed, but *"It wouldn't be fair because people have called it that for so long."*

When 6;2 had been reached she said of dreams, *"I just think of them."* They were known to be a product of the head, particularly of the brain.

Stage 2 of animism was present at 4;7. Stage 3 was not in evidence at any examination. The adult concept was evidenced at 6;2. These records are very far in advance of the progress of the average child (4).

Since artificialism seems not to change in a step-wise manner, but in contrast to the other tendencies, changes with regard to separate objects, it is not easy to trace its decline. Briefly it may be reported that *M*'s artificialism in respect to common objects of nature was reduced at 4;11 and very greatly reduced at 6;2.

J. Discussion

No one can insist that he knows every event in any child's life, and I am not claiming such knowledge here. Therefore the sceptical may insist that the early ideas of *M* were somehow socially transmitted and were not really autogenous. For my own part, I cannot accept this interpretation. As stated earlier, *M* was in the presence of Mrs. Dennis or myself at practically all times during the first three years. During only a short part of this period did we have a servant, and then the servant was occupied with household duties and not with the care of *M*. *M*'s information, therefore, when it did not come from her own observation, came almost exclusively from us.

It need scarcely be said that we did not tell her that one thinks,

wishes, and believes with the mouth, that dreams are in and about the room, that tables and chairs can sense and know and that they are living, that the sun and moon move with the child and may be moved by her, and that mountains, trees, and streams have been made by men. Furthermore, it is hardly conceivable that some person other than ourselves, in some short period, managed to implant such ideas.

To be sure, *M* was not cut off from childhood literature. By 2;9, when she was first questioned in Piaget's manner, she had heard a number of songs, children's poems, and stories. The folk songs were from the collection of Surette, though of course she had not heard the entire group. The poetry familiar to her was from *Mother Goose,* Stevenson's *Child's Garden of Verse,* and Milne's *When We Were Very Young.* No fairy stories had been read to her at that time. The tales with which she was most familiar were *The Three Bears, Little Black Sambo, Little Red Hen,* and *The Gingerbread Boy.*

In some of these stories, speech is put into the mouth of animals. That fact, however, would seem to be irrelevant, since Piaget has asked no questions with regard to what objects can talk.

It may be urged that *The Gingerbread Boy* teaches animism, since in that tale a gingerbread cookie in the shape of a boy comes to life. It should be indicated that the concept of life which is employed in this story is not the concept of the child, since to the child the cookie was alive even before it began to carry out human activities. In other words, the child is much more animistic than is the story.

In the literature known to *M* at 2;9 I am unable to find probable origins for most of the tendencies listed by Piaget and demonstrated in *M*.

Children's literature in its most extravagant form cannot be shown to suggest more than a part of the ideas listed above. Especially is this true of the rhymes which make up the predominant literary fare of *M* under 2;9. It is admitted, however, that the fact that *M* had heard poems and stories is the chief criticism to be urged against my interpretation of her early ideas as autogenous. It would be valuable to examine a young child who had been brought up without contact with children's literature. For the present it seems more reasonable to suppose that the animistic and magical qualities of children's literature appeal to the child because they are familiar

modes of thought rather than to suppose that they come to the child as strange interpretations and yet prevail.

A few comments may be offered concerning *M*'s precocity in regard to these conceptions. It is likely that with many children of 2;9 no answers of the type here presented could be obtained, because of lack of comprehension of the questions. It is certain from Piaget's material (2) and from the results of Russell (4) that mature answers at the six year age level are very rare. *M*'s precocity may be explained on two grounds. In the first place, she spent nearly all of her time with adult companions who talked to her and answered her questions. She was a first born child and there were no other children in the house in which she lived. She spent no time in the care of servants. On several counts her opportunities to gain language comprehension were good. As a second factor may be mentioned her intellectual status. On each of three administrations of the Stanford-Binet between the third and sixth years her *IQ* was 150 or over. Her performance on Piaget's questions is therefore much less remarkable when considered in relation to her mental age at the time of her respective examinations than it is when chronological age alone is taken into account.

The older the child the more difficult it is to keep a record of the day's activities and to limit the social contacts of the child. When one can exhibit all of the features of the early stages of the *Child's Conception of the World* at age 2;9, one's knowledge of the possible antecedents of these answers is more certain than it could be if the conceptions were first evidenced at a later age. After three years of age it becomes more and more difficult to control the child's possible sources of information. The precocity with which *M* gave evidence of the concepts described by Piaget is a strong argument for their autogenesis.

R. Summary

The author's daughter, who was under the close supervision of her parents during the first three years, was asked questions taken from various sections of the *Child's Conception of the World*. The typical stage-one answers reported by Piaget were obtained from this child when she was first examined at two years and nine months of age. The hypothesis that these answers were transmitted by adults is rejected, and it is proposed that they were autogenous, i.e., they developed from the child's own experience and her own reaction tendencies.

The fact that by six years and two months of age this child gave adult answers to Piaget's questions is cited to show the great variations to be observed in the age of disappearance of childish conceptions. Had this child not been examined until she had reached age 6;2, evidence apparently opposed to that of Piaget would have been obtained, whereas the earlier observations show that her development was entirely in agreement with the child's sequences of ideas described by Piaget.

References

1. Dennis, W. Historical notes on child animism. *Psychol. Rev.*, 1938, **45**, 257-266.
2. Piaget, J. The Child's Conception of the World. New York: Harcourt, Brace, 1929. Pp. ix+397.
3. Russell, R. W., & Dennis, W. Studies in animism: I. A standardized procedure for the investigation of animism. *J. Genet. Psychol.*, 1939, **55**, 389-400.
4. Russell, R. W. Studies in animism: II. The development of animism. *J. Genet. Psychol.*, 1940, **56**, 353-366.

Department of Psychology
Louisiana State University
University, Louisiana

The Journal of Genetic Psychology, 1942, **60**, 321-328.

EFFECT OF PERFORMANCE WITHOUT REWARD ON THE RETENTION OF THE MAZE HABIT IN THE WHITE RAT*[1]

Department of Psychology, Columbia University

C. J. Warden and L. Lubow

A. Introduction

A number of investigators have found that when a learning situation is atypical or incomplete in some respect, "latent" learning often occurs. The "latent" learning is reflected in the performance score when the typical or complete situation is presented to the animals later. Warden (7) found, for example, that mere exploration in a simple problem box during the early trials may be thus effective. In fact, a 10-minute period of daily exploration during the first seven days had the same value for the experimental group as one regular trial per day in the control group. Apparently, securing the reward under these conditions tended to decrease the range of exploration to the area in which the latch device was placed when it was introduced on the 8th trial.

Several workers have found that "latent" learning takes place in the maze situation when the usual reward was withheld on the early trials. Blodgett (1) found this to be true when the reward was withheld during the first six daily trials in which exploration was permitted. Haney (4) found evidence that exploration for four days, 18 hours a day, in a multiple *T*-maze was of aid in the later learning of the maze. This advantage was manifest on the first regular trial as well as in the total scores.

Several studies have been made regarding the effect of removal of reward on the retention of the maze habit. Sharp (5) gave white rats 24-60 trials on a square maze without reward after the habit had been established in the usual manner. He found marked evidence of "habit disintegration" during the post-learning period, the amount depending upon the number of trials per day and other factors. Somewhat later, Bruce (2, 3) studied the effect of the

*Received in the Editorial Office on December 5, 1940.

[1]This problem was planned in coöperation. The junior author is responsible for the data and the senior author for the final report.

removal of reward after the maze had been partially learned. He noted that the shift to the non-reward condition, after 10 rewarded trials, caused a disruption in maze performance. Little is known at present regarding the precise nature of such disruption, or habit "disintegration."

The aim of this experiment was to try to throw some light on this problem. The general approach was similar to that of Sharp, in that the maze habit was completely learned before the removal of reward. Our experimental set-up differed from Sharp, however, in two ways: (*a*) the post-learning period ranged from 75 to 135 trials, and hence was longer than the period employed by Sharp; and (*b*) the performance during this period was more continuous, by reason of a special device to be described later.

B. Method and Procedure

The apparatus employed was an 8-unit Warner-Warden maze, modified in certain particulars, as shown in Figure 1.

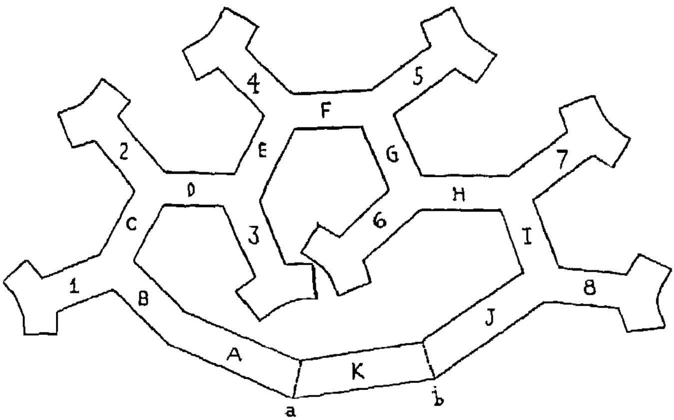

FIGURE 1
Continuous Maze Patterns: Special Modification of the Warner-Warden Maze

The modification consisted of the direct connection (*A, J, K*) between pathway units (*A*) and (*J*) to provide for continuous performance. The middle section (*K*) was used as a common entrance and goal compartment, and was separated from the run way units (*A*) and (*J*) by sliding doors (*a*) and (*b*).

The animal was released from the entrance compartment through door (*a*), and after running through the maze was admitted to the entrance compartment through door (*b*), this door being closed behind the animal to prevent retracings. Since the animal could be started upon the next trial simply by opening the door (*a*), this device eliminated handling of the animal between the trials of a given daily series. This prevented any change in motivation that might arise from the fear of handling, or from the reward of escape from the apparatus. This device tended also to make the series of daily trials more of a continuous performance than would be true in a regular maze situation.

The maze was placed in a dark room with a large indirect light placed directly above it.

A total of 19 male white rats about 70 days old were divided into an experimental group of 13, and a control group of six animals. Both groups were first trained up to the norm of four perfect trials out of five.

During the post-learning period the experimental group was required to run the maze without reward, whereas the control group continued to be rewarded as usual. For the first two days the animals of both groups were given 10 trials daily. This daily test period was reduced to five trials thereafter, because the animals of the experimental group spent so much time in the maze, that a longer series could not be conveniently completed. For the same reason, fewer than five trials per day were given in some cases, since this number was not always completed within the limits of a 20-minute period for each animal. All the animals were tested between 9 A.M. and 1 P.M. Time was recorded with a stop watch and both retracing and blind alley scores were noted for each unit of the maze. The animals in the control group were given a nibble of milk-soaked bread after each trial and fed their regular meal immediately after their daily test period. Each animal of the experimental group (non-rewarded) was fed two hours after the series of daily runs had been completed, the daily food ration consisting of milk-soaked whole wheat bread, and the usual greens.

C. Results: Learning Series

The 19 animals were required to learn the maze up to the norm of four perfect trials out of five before being separated into control and experimental groups. This required an average of 17 trials, the

range being 10 to 30 trials. The two following groups were divided on a basis of speed of learning into (*a*) control (6 animals), and (*b*) experimental (13 animals). The post-learning period, as outlined above, was begun on the day after the learning series was complete.

D. Results: Post-Learning Series

This series differed from the learning conditions for both groups in an increase of the number of trials per day. The number was either 5 or 10 for the control group and as many for the experimental group as the animals would run within these limits. The main difference in the conditions for the control and the experimental group was that the former was rewarded and the latter was not. The results obtained in the two groups will be discussed in order.

1. *Control Group*

The record of errors for the control group are given in Table 1.

TABLE 1
RECORD OF ERRORS OF THE CONTROL GROUP COVERING 135 TRIALS OF THE POST-LEARNING PERIOD

Rat No.	Distribution of errors in the 135 trial series							
	0	1	2	3	4	5	6-10	11-20
1	108	16	4	4	2	1		
2	109	11	6	3	4	2		
3	89	23	12	3	4	1	3	
4	114	16	4	1	0	0		
5	107	21	2	0	22	2	1	1
6	94	26	7	7	0	0		

The actual values for Columns 7 and 8 are as follows: (1) Rat No. 3, 7, 8, 9 errors, (2) Rat No. 5, 10, 16 errors.

Since the animals were rewarded they continued to complete trials throughout the schedule of 135 runs. The animals ran rapidly and consistently, usually completing the trials within 10 seconds.

It might have been expected that this vast amount of overlearning would have led to perfection of the habit. However, this was not the case. As shown in Table 1 errors continued to occur on a considerable number of trials. The number of perfect trials ranged from 66 per cent to 85 per cent for different animals. The typical record was a series of perfect trials—sometimes as many as 10—followed by a run of trials involving errors. As indicated in Table 1,

an animal was likely to make a number of errors, ranging from 1-16, on the same trial. Such relapses sometimes occurred even in the later stages of the post-learning period. In fact more than a third of the total number of errors was made during the last half of the 135 series.

No satisfactory explanation can be offered as to why such an extensive period of practice did not result in complete error elimination. However, it may be suggested that the shift from a single trial to 5 or 10 trials per day may have been effective. The animals obtained more food under the latter conditions. This may have reduced the motivation factor. Moreover, immediate repetition may have favored the stereotyping of errors occurring early in the day's run.

2. *Experimental Group*

Since this group was not rewarded, the animals tended to hesitate or explore, and took an unusually long time to reach the food box. It was necessary, therefore, to set a time limit for each day's testing. Each animal was given as many trials as he would run in 20 minutes, the trials being limited to 10 runs. This ranged from zero, when an animal refused to run, up to the maximum of 10 trials. As a rule, animals of this group ran about five trials within this 20-minute period. The total number ranged from 75 to 135 trials under these conditions.

As might be expected the animals in the experimental group made errors on more trials than did the animals in the control group. Only 39 per cent of the trials were errorless, as compared with 76 per cent in the control group. Moreover, the number of errors per trial was very much greater for the experimental group, as can be noted by comparing the values of Tables 1 and 2. The poorest for the control group was 16 errors on one trial, while that for the experimental group was 74 errors. The number of trials on which more than five errors were made was markedly greater for the experimental group. Nevertheless, 4 of the 13 animals (Nos. 1, 7, 11, 13) made rather remarkable records in spite of the absence of reward. The scores of errorless trials for these animals ranges from 45.6 to 51.6 per cent. With two exceptions the animals showed runs of four perfect trials out of five from two to ten times during the post-learning period. Rat No. 5, after making 74 errors on trial No. 12, was able to make a score of four perfect trials out of

TABLE 2

RECORD OF ERRORS OF THE EXPERIMENTAL GROUP COVERING THE POST-LEARNING PERIOD (75-135 TRIALS)

Rat No.	No. trials	Distribution of erors by trials 0	1	2	3	4	5	6-10	11-20	21-30	31-40	41-74
1	125	57	28	15	8	4	6	3	2	2		
2	135	44	20	18	16	11	6	12	8			
3	112	48	23	10	6	10	3	6	4	2		
4	75	34	12	6	8	4	2	6	3			
5	87	33	16	9	9	1	0	7	3	4	3	2
6	135	44	31	20	13	11	8	4	4			
7	123	63	25	9	6	4	3	8	5			
8	84	19	13	17	6	8	5	10	4	1	1	
9	82	32	24	2	8	4	4	7	3			
10	84	29	18	11	3	3	3	9	7	1		
11	118	57	5	7	7	5	5	10	14	7	1	
12	85	14	19	8	9	2	3	12	13	2	3	
13	124	64	16	14	9	3	3	7	6	1	1	

The actual values of Column 13 are as follows: Rat No. 5, 52, 74 errors.

five no less than five times. On the whole, more errors were made by the experimental group during the first half of the series than during the last half.

These results would seem to call into question the concept of "habit disintegration" in so far as it implies memory loss. It is true, of course, that the habit is temporarily disturbed by the removal of reward, but a similar disturbance, although less severe, occurs during post-learning practice when the incentive is available. Only the difference between results of the control and experimental groups can be properly ascribed to removal of reward. In neither case can it be said that the habit had shown permanent disintegration. This is clear from the fact that errorless runs occurred often in fairly long sequences during the post-learning period. As a matter of fact, the percentage of errorless runs was more than half as great for the experimental group (39 per cent) as for the control group (76 per cent).

These facts indicate that the habit pattern is still retained and can be evoked from time to time. This conclusion is in agreement with the results reported by Tsai (6) and others, showing that maze habits are retained for several months in the white rat with practically no loss. Thus, the phenomenon of "habit disintegration," reported by Sharp and in this paper, is concerned with a disturbance in performance rather than in memory. The removal of reward has merely a temporary influence on performance rather than a permanent effect on retention.

E. Summary

A group of 19 white rats were trained on a Warner-Warden maze until a norm of four perfect trials out of five was reached. The animals were then divided into a control group (6 animals), and an experimental group (13 animals), on the basis of their learning record. Both groups were given a long series of post-learning trials, ranging from 75 to 135. The control group was rewarded after each trial and the experimental group was run without incentive.

The following results and conclusions are supported by an analysis of the data:

1. The extensive post-learning practice with an incentive present, under our conditions, did not lead to complete error elimination (control group). Only 76 per cent of the trials were errorless, and a number of errors might be made on a trial even late in the post-learning series.

2. The removal of incentive (experimental group) merely increased the tendency to make errors, the number of errorless trials being only 39 per cent. The number of errors per trial was considerably increased over that of the control group.

3. The concept of "habit disintegration" is interpreted in terms of a temporary disturbance of performance rather than as a permanent disintegration of habit.

References

1. Blodgett, H. C. The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 113-134.
2. Bruce, R. H. The effect of removal of reward on the maze performance of rats. *Univ. Calif. Publ. Psychol.*, 1930, **4**, 203-214.
3. ———. The effect of removal of reward on the maze performance of rats: II and III. *Univ. Calif. Publ. Psychol.*, 1932, **6**, 67-82.
4. Haney, G. W. The effects of familiarity on maze performance of albino rats. *Univ. Calif. Publ. Psychol.*, 1931, **4**, 319-333.
5. Sharp, W. L. Disintegrative effects of continuous running and removal of the food incentive upon a maze habit of albino rats. *J. Comp. Psychol.*, 1929, **9**, 405-423.
6. Tsai, C. A comparative study of retention curves for motor habits. *Comp. Psychol. Monog.*, 1924, **2**, No. 11.
7. Warden, C. J. The value of the preliminary period of feeding in the problem box. *J. Comp. Psychol.*, 1925, **5**, 365-372.

Department of Psychology
Columbia University
New York City

The Journal of Genetic Psychology, 1942, 60, 329-335.

STUDIES IN ANIMISM: V. ANIMISM IN OLDER CHILDREN

Department of Psychology, Michigan State College

ROGER W. RUSSELL

A. INTRODUCTION

The development of an interest on the part of the author in the experimental study of child animism has not been one solely of the investigation of children's ideas, but also has been stimulated greatly by the desire to investigate as completely as possible the ontogenetic manifestations of higher thought processes. It was toward this end that studies culminating in the reports by Russell (2, 3), Russell and Dennis (4), and Russell, Dennis, and Ash (5), were initiated. The present paper furnishes the final step in the study by reporting the results of an investigation in which the development of the animistic concept is carried through to a point where a majority of the subjects have attained the final or adult stage of concept.

B. METHOD

The present procedure differs from that employed in previous investigations of child animism (4) in that it is adapted to comply with conditions of group testing. The subjects were gathered together in a single room under the direction of the examiner. The questions concerning the animate or inanimate nature of a series of objects, almost exactly similar to the list employed by Russell and Dennis (4), were mimeographed on single sheets of paper. These sheets gave directions as to what should be done with the test items and furnished sufficient space for the subjects to write down the reasons for their classifications of the items as animate or inanimate. Table 1 presents a list of the test items. The directions to the subjects were as follows:

1. Fill in the blanks at the top of the page. When you have finished, hold up your pencil.
2. Look at the list of names below. If an object is *living*,

*Received in the Editorial Office on December 10, 1940.

TABLE 1
List of Objects in Serial Position of Presentation

1. Knife	6. Dish	11. Wind	16. Tree
2. Mirror	7. Watch	12. Bicycle	17. Flower
3. Button	8. River	13. Dog	18. Grass
4. Comb	9. Clouds	14. Bird	
5. Chair	10. Wagon	15. Bug	

or is *sometimes living*, write "yes" on the short line to the right, and then write on the long line the reason why you think it is so. If an object is *never living*, write "no" on the short line, and then tell why it is never living.

3. This is not a school test. Just write down what *you yourself* think. You do not need to hurry. Be sure to fill in every line.

4. When you have finished, hold up your pencil.

The examiner[1] followed the same procedure at each administration of the scale. No questions were answered that might in any way have suggested an answer and no examples were ever given. Collaboration among the subjects was prohibited.

The type of responses obtained were easily classifiable into one or the other of the stages characteristic of the development of the animistic concept. Due to the method of stating in the directions the basis for the classification there appeared in the answers instances in which certain objects were considered as once having been living and therefore classified as animate. This occurred most frequently in the case of the knife which to the particular child was considered as having a wooden handle, and in the case of the wagon, which was sometimes considered as being made of wood. Since the wood had once been living, the object was classified as animate in view of the statement in the directions that an object which "is sometimes living" must be so classified. These instances were relatively few and easily discerned when they did occur.

In general the classification of responses into the various stages of animism was as easily accomplished as in the case of responses obtained by individual examination. Classification of responses by trained examiners showed an agreement of 95 per cent. The difficulties that contributed to the disagreement in the other five per cent

[1]The writer is greatly indebted to Miss Josephine W. Gregg for invaluable assistance in the administration of the scale.

arose from the inadequacy of the data in these instances. This latter fact is the only apparent handicap to the use with older children of such a group examining procedure as has been employed in the present study.

C. Subjects

The present study reports results obtained from the group examination of 611 children.[2] Of these 162 were enrolled in the Cherry Creek High School in Cherry Creek, New York; 182 were in the Orchard Park High School in Orchard Park, New York; and 267 were students at the Wilson Junior High School in Hamilton, Ohio. Cherry Creek is a small agricultural village in upper New York state, while Orchard Park is a residential suburb of Buffalo. Hamilton, Ohio, is a manufacturing city of more than 50,000 population.

These three samplings represent a wide divergence of environments[3] and it is interesting in this regard that there are very normal distributions of intelligence quotients for each of the individual groups. The median *IQ* for the total distribution is 101.35 with a range from 60 to 147. Mental ages were obtained by means of the

TABLE 2

DISTRIBUTION OF SCORES ON THE OTIS GAMMA EXAMINATIONS FOR THE COMPOSITE EXPERIMENTAL GROUPS

IQ	Frequency	
145-149	1	
140-144	2	
135-139	3	
130-134	9	
125-129	12	
120-124	20	
115-119	28	
110-114	69	
105-109	74	
100-104	128	Median—101.35
95-99	77	
90-94	76	
85-89	42	
80-84	33	
75-79	18	
70-74	14	
65-69	2	
60-64	3	Total—611

[2]Complete data are not available for a few of these subjects.
[3]Note conclusions concerning the effect of environment in Russell (2).

Otis Beta and Gamma Examinations. All testing was conducted by members of the school systems trained for this work. Table 2 presents the composite test results. The range in chronological age was from 8 to 20 years and the range in grade from 5 to 12.

D. Results

1. *Classification into Stages*

Only 11 of the 611 subjects had systematic distinctions between the animate and the inanimate which were not in agreement with any of the four animistic stages. This means that, of all the subjects other than the one subject who was in the *No Concept Stage,* only 1.8 per cent did not definitely fall into one or other of the designated stages. These were classified, as in previous studies (2, 4, 5), in a Special Concept Stage.

2. *Progression of Stages*

Figure 1 and Tables 3, and 4 show the percentage of cases in each stage of concept for two year periods from mental and chronological ages of 8 years 9 months to 19 years 11 months.

TABLE 3
PER CENT OF CASES IN EACH STAGE OF CONCEPT FOR TWO YEAR *MA* PERIODS

MA	*NC*	1	2	3	4	*SC*	*N*
8^{0}-9^{11}	0.0	0.0	20.0	10.0	40.0	30.0	10
10^{0}-11^{11}	1.6	3.0	17.2	34.4	43.7	0.0	64
12^{0}-13^{11}	0.0	0.0	10.8	26.9	61.1	1.2	167
14^{0}-15^{11}	0.0	1.0	7.8	21.8	69.6	0.5	204
16^{0}-17^{11}	0.0	0.0	0.0	10.1	86.9	4.0	99
18^{0}-19^{11}11	0.0	0.0	0.0	3.0	90.9	6.1	33

TABLE 4
PER CENT OF CASES IN EACH STAGE OF CONCEPT FOR TWO YEAR *CA* PERIODS

CA	*NC*	1	2	3	4	*SC*	*N*
8^{0}-9^{11}	0.0	0.0	33.0	0.0	33.0	33.0	3
10^{0}-11^{11}	0.0	0.0	20.0	28.0	52 0	0.0	25
12^{0}-13^{11}	0.5	0.5	11.6	27.5	59.9	0.0	207
14^{0}-15^{11}	0.0	1.1	7.3	20.2	69.5	1.1	262
16^{0}-17^{11}	0.0	0.0	2.2	13.0	80.4	4.3	92
18^{0}-19^{11}	0.0	0.0	0.0	7.7	92.3	0.0	13

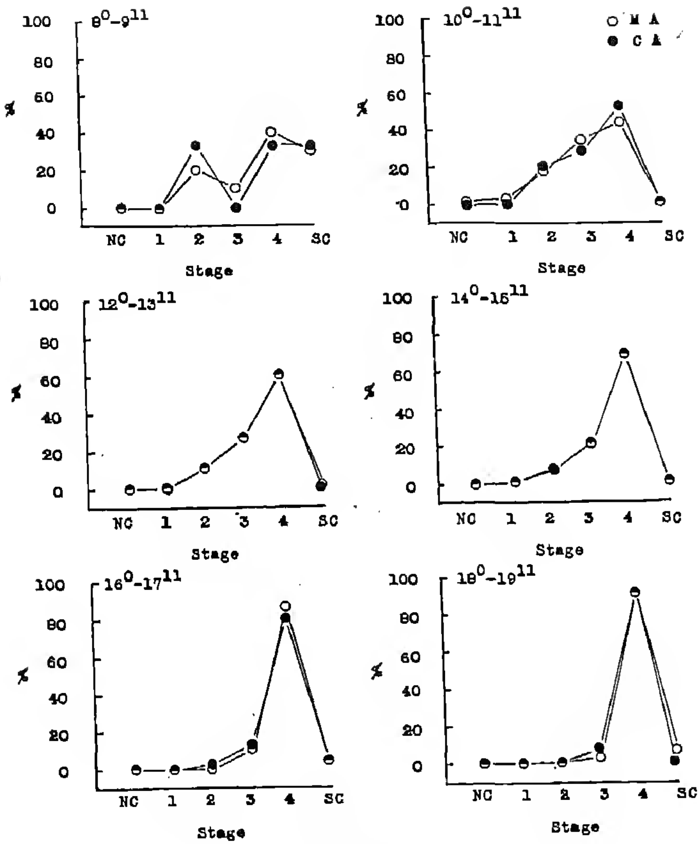

FIGURE 1

PER CENT OF CASES IN EACH STAGE OF CONCEPT FOR TWO-YEAR PERIODS

From Figure 1 it is apparent that the percentage of cases in the adult stage or Stage 4 increases with both mental and chronological age. In the 18^{0}-19^{11} age group over 90 per cent of the cases fall in Stage 4 classification. As the percentage of cases increases at the higher levels the expected decrease in percentages at the lower stages occurs until cases at the *NC* stage, at Stage 1 and at Stage 2 have disappeared entirely in the 18^{0}-19^{11} mental and chronological age groups.

3. *Correlation of Stage with MA and CA*

The correlation between stage and animistic concept and mental age was found to be 0.49, while the correlation between stage and chronological age was 0.44. Since these correlations are expressed in terms of Coefficients of Mean Square Contingency derived from a 6 x 6 contingency table the correlations could not be expected to exceed 0.91. However, despite this fact, the correlations still express the existence of relationships far too slight for accurate individual predictions. To be sure, there is indicated some relationship but, in this case, that relationship is even less than the similar ones existing in the development of the animistic concept among younger children.

4. *Correlation of Stage with Grade in School*

The possibility of training in school forming the fundamental basis for the ontogenesis of animism gained little support from the correlations between school grade and stage of concept. The coefficient of Mean Square Contingency in this instance proved to be only 0.40.

E. Conclusions

A consideration of the present results seems to warrant the following conclusions:

1. It is possible to classify older children into the stages of animism found to exist among younger children. This is substantiated by the fact that 98 per cent of the subjects studied were readily classifiable into one or other of the stages of animism.

2. There is a definite increase in the percentage of individuals at the adult stage (Stage 4) of concept accompanying progressive increases in mental and chronological age. In the 18^{0}-19^{11} mental and chronological age levels over 90 per cent of the cases fall in Stage 4, while no cases whatsoever appear at the *NC*, Stage 1, or Stage 2 levels. This completes the ontogenetic picture of animism which started with 100 per cent of the subjects in the No Concept Stage at early nursery school levels.

3. The stages of animism are equally related to both mental and chronological age, the Coefficients of Mean Square Contingency for the group being 0.49 and 0.44 respectively.

4. The Coefficient of Mean Square Contingency for stage of animism and grade in school is 0.40.

References

1. Dennis, W., & Russell, R. W. Piaget's questions applied to Zuni children. *Child Devel.*, 1940, **11**, 181-187.
2. Russell, R. W. Studies in animism: II. The development of animism. *J. Genet. Psychol.*, 1940, **56**, 353-366.
3. ———. Studies in animism: IV. An investigation of concepts allied to animism. *J. Genet. Psychol.*, 1940, **57**, 83-91.
4. Russell, R. W., & Dennis, W. Studies in animism: I. A standardized procedure for the investigation of animism. *J. Genet. Psychol.*, 1939, **55**, 389-400.
5. Russell, R. W., Dennis, W. & Ash, F. E. Studies in animism: III. Animism in feeble-minded subjects. *J. Genet. Psychol.*, 1940, **57**, 57-63.

Department of Psychology
Michigan State College
East Lansing, Michigan

The Journal of Genetic Psychology, 1942, 60, 337-357.

THE QUESTIONING ACTIVITY OF CHILDREN*[1]

Department of Psychology, University of Wisconsin

GEORGE L. FAHEY

The present article is a summary of the literature dealing with the questioning activity of children with particular emphasis upon those inquiries that are made in the classroom. A functional organization has been attempted which is based upon the types of problems attacked. Only where chronology had some particular significance has it been stressed.

A. THE CONSTITUENT ELEMENTS OF A QUESTION

The literature reveals little of value bearing upon the nature of the question. Fundamental assumptions are made about the meaning of the term, but there is a lack of any satisfactory definition. Earlier writers considered the question as an expression of an "instinct of curiosity." Bain (1), Bohannon (3), Sully (45), and Compayre (10), writing between 1879 and 1902, adopted this point of view but added that excessive questioning might not indicate any real curiosity. To Sully, a child's questions might be motivated by a "desire for order and connectedness, leading to anthropomorphism, teleological explanation, and a deeper sense of perplexity, mystery, and contradiction." This suggests that children's thoughts are penetrated with the idea of purpose and use. Bain added that incessant questioning probably serves to display ego and to create a satisfying state of excitement with the questioner as the center of attention.

More recent writers have, in keeping with the times, avoided reference to instinctive origins. Wallon (50) believed that questions marked a break in mental operations due to some strange intervention and served to prepare the questioner for the response. He concurred with the earlier writers in the belief that excessive questioning may be pathological (sic) in nature. Stern (42) contended

*Received in the Editorial Office on January 7, 1941.

[1]This study is one of a series of investigations supported in part by grants from the Research Committee of the University of Wisconsin Graduate School and completed under the general supervision of Stephen M. Corey of that institution.

that the questioning activity of children under the age of four years was the result, first, of attempts to get some justification of commands, and, second, the desire to know. Johnson (28) attributed the asking of many questions to a desire for social contact. Hollingsworth (26) defended these multiple origins: first, genuine curiosity; second, the desire on the part of the child to check his own generalization against those of one better informed, and, third, getting and holding the attention of others. Griffiths (22) illustrated curiosity by reference to the incessant questioning of young children and defined "curiosity" as "an activity attitude very much like interest."

The fundamental nature of the question has been "almost totally ignored by the logicians" (Cohen, 9). In Cohen's judgment the contentions that a question is only the beginnings of thought leading to an end;—or that it is no more than a request for information; or that it is simply an ambiguous assertion, are superficial. It is none of these, but something more,—a "logical entity"; a propositional function set in words with a variable (who, which, what, why, and so forth) substituted for some constant. Only in form does *What is four plus five?* differ from $x = 4 + 5$. Both contain a variable and have a constant as an answer; hence, "questions are variables whose values are answers." The term "question" should be restricted to the thought content and "interrogation" used as the verbal symbol of the inquiry.

Claparede (8) ventured this definition of the term as follows:

> A question is the conscious realization of a problem or of the difficulty in solving it; i.e., of the direction in which one must seek for its solution. The nature of this question will determine the whole course of subsequent research. Thus the function of the question is quite clear; it is an incitement to mental activity, *in a certain* direction in view of readjustment.

Symonds (46) discussed questions and problems almost synonymously although a slightly more restricted concept may have been implied by "question." A problem may produce questions and apparently must do so if it is to be solved effectively.[2] Whenever a question occurs it implies that two or more objects in a field are kept apart by some gap in the set of relationships connecting them. The possibility of relationship must be suggested, however, before the

[2]It does not seem unlikely that a question may also give rise to problems.

question takes form. Wrinkle and Armentrout (54), on the other hand, distinguished between questions and problems by considering questions as representing potential problems,—questions which have become matters of personal concern to the individual.

According to Woodworth (53) the question is "interest in a definite form," originating in curiosity arising out of some particular experience. It serves to prepare the individual for reception of the answer and the exclusion of that which is irrelevant.

Discussing "why" questions, Isaacs (27) emphasized the aspects of the situation giving rise to "whys" as being one of a sudden clash between present and past experience. Several shades of meaning exist for "why" but all are related originally to some state of un-expectedness, anti-expectedness, or confusion of expectation.

The writer (19) has defined a question as a "verbal expression of a problematical situation existing in the mind of the questioner," with the added comment that problematical was to be interpreted broadly enough to include realization of lack of specific information, awareness of gaps in relationships, or a consciousness of a conflict between present and past experience. The purpose of the question in terms of this definition would be the finding of a means of supplying the lack of information closing the gap, or resolving the conflict. This defintion obviously does not consider those inquiries made to gain time or attract attention as being questions in any psychological sense.

B. The Pedagogical Importance of Questioning Activity

The theoretical importance of questioning activity in the processes of thinking and learning has been frequently stressed in the pedagogical and psychological literature. Few texts dealing with educational methods or theories of learning, particularly those of recent years, fail to stress those aspects of the learning situation which involve the solution of a problem or the determination of an answer to a question. The following quotations are typical. A much greater number could be accumulated.

> " . . . instruction, then, in subject matter that does not fit into any problem already stirring in the subject's own experience, or that is not presented in such a way as to arouse a problem, is worse than useless for educational purposes" (15, p. 199).

The question "reinforces the effect of certain objects and incidently inhibits the effect of other objects, for observation sharpened by a question is keen only for the answer to that question and neglectful of whatever is irrelevant to it" (53, p. 122).

"Another good method is to have the class prepared to ask definite questions concerning the work of the previous day's assignment. . . . Definite questions show that a pupil has been thinking about his work and in addition give the teacher a cue as to just what the difficulty is" (24, p. 172).

"In any region that is new, strange, and unfamiliar one can plow his way by means of questions. Wherever the field begins to become vague, indistinct, and fuzzy at the edge, one can build up the structure or the organization of relationships by appropriate questions which will clarify concepts, place terms in proper sequence or classification, organize hierarchies of values and the like. The question, then, specifies the gap in relationships and directs the active search toward closing the gap. The question is helpful in directing the attention and thus aids in analysis" (46, pp. 47-48).

"The question is a natural expression of the thinking mind. The teacher who does not receive a number of unsolicited questions should seriously examine his methods" (52, p. 489).

"The one seeking knowledge should have first claim upon the question. . . . The problem of encouraging pupils to ask questions should receive the teacher's consideration while planning his classroom activities" (49, p. 272).

C. The Questioning Activity of Young Children

A large number of the observational reports of questioning activity have been concerned with the acquisition of language by young children and the time of initial appearance of various types of questions. In the order given, Sully (45), Preyer (37), Trettien (48), Stern (42), Norsworthy and Whitley (33), and Kenyeres (29) have reported data on the appearance of interrogative words. Their methods preclude any exact comparisons but in general the agreement is marked. Sully (45) reported "what," "who," "when," "where," and "how old" as appearing "earliest," and "why's" as "beginning at two but most frequent between three and four years." Preyer (37) reported "ou" (where) at 28 months, "pourquoi" (why) at 34 months, and "quand" (when) at 36 months. Trettien (49) observed that "questions of fact, substance, action, and place"

—the "what" and "where" type—were earliest and that "why's" appeared later. Stern (42) found questions about names to appear after the eighteenth month, questions involving "when" and "why" during the third and fourth years. Norsworthy and Whitley (33) placed "what" between the eighteenth and twenty-fourth months and "when," "what is," and "why," in the order given, about the third year. Kenyeres (30) reported questions about the qualities of objects at the end of the second year; "when" about the third year, and "why" as first appearing at one year and nine months, quite frequent at two and one-half and very frequent at seven years.

Wallon (50) asserted in the same connection that the earliest questions may be interrogative only by virtue of the attitude of the child when speaking. He objected to Preyer's (37) order of the chronological appearance of question forms on the grounds that the inquiries represented may have been an order of desires rather than one of time of acquisition of a form of speech. His objection is undoubtedly pertinent although the commonly observed order of naming, location, time, and reason is suggestive of the logical acquisition of more complex concepts which might be expected to appear in some such order as the child grows older.

The period designated as the "questioning age" in child life has been defined thusly: as occurring during the fourth year of life (Sully, 45); as beginning in the third year but most fully developed in the fourth (Trettien, 48); as between the third and sixth years (Hollingsworth, 26) and as at two periods; first, that of naming, reached in the second half of the second year, and, second, that of "when" and "why" reached in the third and fourth years (Stern, 42).

These observational and anecdotal reports do not lead to any sweeping generalizations. No one has actually defined the various interrogative types, then made observations of their order of appearance among various age groups, and then reported central tendencies. The writer (19) suggested elsewhere that there is a possible relationship between a hierarchy in terms of the thought implications of questions and the age at which various question forms first appear. The establishment of such a hierarchy must, however, await the development of standards for assigning thought values to questions and a less disputed chronological order of appearance of question forms.

Several writers have reported objective evidence of the extent of questioning activity among children as measured by the per cent of all conversation which was devoted to questioning. Trettien (48) reported one subject between the ages of three and four years who devoted 22 per cent of all words used in conversation to questioning. The Brandenburgs (6) observed their daughter to use questions at the rate of 31 per hour (18 per cent of all her conversation) at the age of 38 months, while at the age of 52 months, she asked 33 questions per hour (20 per cent of all her conversation). Nice (32) observed one subject at the age of 47 months who devoted only 11.2 per cent of his conversation to questions. Boyd (5) made observations of his daughter at yearly intervals from the age of two through the eighth year. At two years, questions made up two per cent of all her conversation, at three years 28 per cent, and, for the next seven years 21 per cent. Piaget (35) observed two boys of six years of age over a period of 10 months and found one to devote 17 per cent, and the other 13 per cent of all conversation to the asking of questions.

From these reported observations it is apparent that questioning activity is a major form of oral speech among children of pre-school age. It is made clear in the studies of the questioning of children of school age reported below that the frequency of questions asked in the classroom, tends to decrease as children grow older. That this reduction in the amount of questioning activity is a regrettable outcome of experiences in the home and school which tend to inhibit questioning has been asserted by several writers (16, pp. 44-45; 33, p. 272; 24, pp. 194-195; 55, p. 106; 46, pp. 48-49).

McCarthy (31) compared nursery school children from homes of high and low socio-economic status and noted that between the ages of 36 and 54 months questions comprised seven per cent of the conversation of children from homes of low socio-economic status and 14.4 per cent of the conversation of children from homes high on the same scale. She reported further that at 18 months only four-tenths of one per cent of the conversation of boys and four and seven-tenths per cent of the conversation of girls was given over to questioning. At 24 months, the "questioning" percentages were, for the boys, two and eight-tenths per cent, and for the girls, five and three-tenths per cent. Smith (39) found the relative amount of conversation devoted to questioning to increase up to the age of

five years which was the maximum age limit of her nursery-school subjects. Davis (14) reported that the length of time to collect 50 consecutive questions from various children between the ages of 3 and 12 years ranged from 30 minutes to 14 hours and five minutes. Part of this difference may have resulted from inconsistencies in the method of recording questions.

In 1926 Piaget (35) published a comprehensive two-volume study of the language and thought of children. A large portion of the first volume was devoted to questioning and especially to the questions asked by a six-year old boy and recorded over a period of 10 months at daily intervals of two hours each. Piaget was chiefly interested in "why" questions and for a part of the period recorded inquiries employing that interrogative only. Most important of Piaget's treatment of questions for subsequent research was the functional classification of questions which he devised and reported in such detail that it could be used by other observers. References will be made below to details of the classification system. Piaget was much impressed by ego-centricism in child thought and arrived at an index figure thereof. Inflected forms of questions were not included. The conversation of two boys was recorded, but only that of one was used in the analysis of questioning activity.

Boyd (5) recorded from his daughter's speech, 1,250 unselected sentences for each year between her second and eighth birthdays. The percentage of questions which he found has been reported above. Comparing the records of the child's conversation with that of adults in modern fiction, Boyd found first, that questions are more common in childish speech than in the speech of adults, and, second, that women ask more questions than men, but variation in the number of questions asked is greater among men. The assumption that genuine individuals correspond to fictitious characters in the extent of their questioning is essential to acceptance of these comparisons.

The first statistical data bearing upon differences among individuals in the amount and character of their questioning activity were reported by Davis (14) in 1932. She analyzed carefully 3,650 questions asked by 73 children between the ages of 3 and 12 years. The Piaget system of classification was used to discriminate question types and the reliability of the method was reported. These coefficients of reliability were 0.88 between two classifications by the

author and 0.77 between the author's classification and that of a colleague. Much dependence was placed on the basic data which were secured by having the mothers of the children record 50 consecutive inquiries for each child.

Davis found significant differences between individuals and groups of individuals as follows: boys asked questions at a faster rate than girls; the number of words per question increased with age for both sexes but questions by girls were longer at all ages; boys began questions more frequently with an interrogative word; boys asked more "why" questions and questions of casual explanation than did girls; girls asked more questions of social relations than did boys (this type of question is a variation of the Piaget classification system); boys tended to ask more questions of definition than did girls, and, as chronological age increased, children of both sexes asked significantly more questions about human actions and intentions and of corroboration and approbation. Eighty-six per cent of the questions were asked of adults and 88 per cent were stimulated by the immediate environment. Compared with 300 adult questions taken from legal testimony and 200 taken from modern fiction, the only difference observed was the greater length of the adult questions.

McCarthy (31, see above) also used the Piaget classification and reported comparisons between the questions of nursery school children from homes of high and low socio-economic status. Smith (39) reported a carefully controlled study, involving 219 pre-school children aged 18 to 72 months. The data analyzed consisted of 305 complete conversation records. No significant differences were observed between the sexes in terms of the total number of questions asked for the period, although at the age of two years the girls asked nearly twice as many questions as the boys. Children were found to ask significantly more questions of adults than of other children, but the tendency to do so decreased with age. There was a marked tendency for the total number of questions asked per child to increase with increase in chronological age. "What" and "where" questions decreased with age; "who," "whose," and "which" did not change appreciably; "how," "when," and "why," increased regularly and significantly from year to year.

D. Scientific Investigations of the Questions of Children of School Age

Despite frequent assertions of the value of pupil questioning in the

learning process and the comparative abundance of studies dealing with young children, the number of scientific reports involving questioning in the classroom is very small. Most of those that have been made deal with children in the elementary grades and have as their primary object the selection of textual or curricular materials. A few deal with the thought implications of pupils' questions, study habits implied by pupils' questions, or pupils' questions as an auxiliary factor in some larger situation.

In 1912, Mau (30) and Downing (17) attacked the problem of the interest implications of children's questions. Each worked independently and with different techniques. Mau set out to discover children's interests in nature materials by bringing into kindergarten and primary-grade classrooms certain objects and recording the spontaneous questions which were asked. The questions were classified according to the types of interests implied. She found boys as a group and as individuals asked more questions than did girls. Nearly all questions were concerned with the names and activities of the objects observed. Downing drew his data from the question and answer department of *St. Nicholas Magazine* and classified the questions according to interests reflected. By nature of the data Downing collected, individual differences were not observable.

Finley (20) collected over 8,000 questions from 1,713 children of grades three to eight and grouped them according to types of interests implied by the questions. He used the simple but effectual technique of bringing into the classroom a species of salamander known as a "mud-puppy" and recording the spontaneous questions asked about the animal. The request for a name was almost universal; attempts at classification were rare below the fourth grade and infrequent above it; there was a marked falling off of teleological questions in the upper grades and an increase in the number of questions dealing with structure. Forty days after the experimental showing of the animal, the children were asked to write a theme about "the mud-puppy". From the facts brought out in these themes Finley concluded that the children remembered best the points about which they asked the most questions.

Palmer (34) investigated children's interests in nature by securing from rural school teachers some 500 questions which had been asked by their pupils. These questions were classified according to the objects to which they referred. He observed that the number of absurd questions were surprisingly small,—a possible consequence of

the method of collection,—and stated that "everything points to the conclusion that the questions given were such as might have arisen from the inability to interpret observations which had been made more or less deliberately."

Dale (13) collected from high school pupils over 27,000 questions related to health in order to determine to what extent current textbooks serve to provide answers to children's health problems. He then compared the interests reflected in these questions with the content of health textbooks. Questions were gathered by the use of a questionnaire.

In addition to the study described above by Dale, several other investigations have been made of desirable curricular and textbook content in terms of children's interests. Washburne (51) visited upper elementary grade classrooms periodically during one year and offered to answer questions relating to general science. He thus collected over 2,000 questions, classified them according to the objects about which they were asked, and used the results as a partial criterion for the selection of material for a general science textbook. Berry (2) secured from teachers of home economics 1,000 questions which had been asked by pupils. These questions were classified according to the topics about which they were asked and used as a basis for the selection of content for a course in household chemistry.

Pollock (36) sent a questionnaire to eighth grade pupils in 13 city schools and received 3,500 questions which pupils would like to have answered in a course in general science. Results were tabulated on the basis of nouns used and these results were compared with topics treated in general science textbooks. He found no differences between Grade 8-A (high) and Grade 8-B (low) in special interests, but the interest span of the higher group was seven and seven-tenths per cent greater than that of the lower. The range of interests for the girls was 10 per cent greater than for the boys.

In the same year, Curtis (12) sent a questionnaire modeled after Pollock's to several hundred junior high school pupils and to their parents and received some 3,300 pupil questions and 3,332 adult questions on scientific problems. The only significant results involved the interests revealed. No startling differences in grade or sex were outstanding, although technical processes were of more interest to boys and men than to girls and women.

Helseth (25) investigated experimentally the possibility of im-

proving the thinking of seventh and eighth grade pupils in United States history by encouraging them to ask questions in the classroom. She used 16 cases,—two boys and 14 girls,—enrolled in a teacher training school. Apparently as the result of an emphasis on the problem-solving method during the course of a year, pupils greatly increased their abilities to find questions to ask, to ask longer and more thoughtful questions, to participate more actively in the classroom procedures, to solve the problems which they set for themselves, to organize the materials which they collected, and to think seriously about their methods of work. In addition to these gains, the pupils exceeded the norms on standardized tests in United States history, despite the fact that one-third of the year's work had been on problems not covered by the tests. In May the pupils were doing 86 per cent of the talking in the classroom, while in September they had been doing but 40 per cent.

An informal but interesting study of the comparative merits of questions asked by teachers and by pupils was made by Storm (44) in 1928. He visited fourth and fifth grade classrooms and recorded 100 questions which he heard asked; 50 pupil questions and 50 teacher questions. These questions were assembled in random order and duplicated copies were distributed to 40 teachers in a teachers' meeting with the instructions that the "best" questions on each page be checked. The median number of teacher questions checked was 17, while the median number of pupil questions checked was 27. Only two teachers checked more teacher questions than pupil questions. One teacher checked 30 pupil questions and only six teacher questions. Storm concluded that, "If pupils ask as good questions as do the teachers, the pupils certainly ought to be the ones to ask the questions, because almost all will agree that the person who is doing real thinking is the one who is asking himself or someone else questions."

In 1928, Gatto (21) sought to appraise the thinking done by 408 pupils in elementary courses in history, geography, nature study, arithmetic, and literature. Four thousand, one hundred and eight questions were collected in three ways: (*a*) in 100 class lessons, where teachers made records of the questions asked by their pupils; (*b*) in 40 other class lessons, where the writer recorded the questions asked, and (*c*) in 20 "case-studies," in which the pupils were presented with short written narratives, allowed to read them and ask any questions they cared to. These questions were recorded by the interviewer.

Comparisons of the questions recorded by the three methods of collection showed much similarity. Pupil questions were compared with study questions in textbooks and the conclusion reached that "pupils ask questions of types found in their textbooks with few exceptions, and therefore demand, by their questions the same type of study activity which may be found in questions demanded by experts."

In this same study questions were classified according to a 22 category system in terms of the type of study activity implied by the questions. Over 50 per cent were questions of memory for specific facts previously presented in class; 21 per cent were for specific facts not previously presented; 6 per cent were questions implying causal reasoning. All types of questions occurred with approximately equal frequency in all subjects, in all grades, at all the represented levels of chronological ages and intelligence. No statistical indexes other than percentages were presented. The questions of pupils appear to have been considered as very like those of teachers because of similarity in form. The difference in purpose, however, suggests to the present writer that they are not synonymous in meaning even though the words may be identical.

Corey and Fahey (11) investigated the possibility of estimating the complexity of a pupil's thought from the questions which he asks in the classroom. They submitted to 12 judges a list of 200 mimeographed questions asked by 10 seventh-grade pupils in a class in science. The judges were instructed to "check those questions which imply thinking on the higher levels." Intercorrelations of the checks of the judges averaged about +.50 with a range of from —.13 to +.97. Intercorrelations among groups of judges average approximately +.85. Whether or not mental complexity was the point of agreement, the judges did agree markedly on something. A correlation of +.50 was found between the number of times a question was checked as implying mental complexity and the number of words in the question.

The writer (19) had trained reporters record all the classroom questions asked by pupils and teachers in six high school classes throughout an academic year. The number of questions asked per pupil bore a slight but persistent positive relationship to desirable interests, intelligence, academic achievement, reading ability, reading appreciation, accuracy, and consistency in thinking in problematical situations. The relationships were commonly curvilinear in nature; pupils having extremely unfavorable scores on the various measures

tended to ask a markedly greater number of questions, but some pupils at all levels of favorable scores asked almost no questions. The types of questions asked, according to the Piaget classification system, were only slightly related to the other pupil variables.

The conclusion to this investigation was as follows:

> The study indicates that the classroom questioning activity of high school pupils is very limited in quantity, diversified in type, and only slightly related to measures of intelligence, interests, reading abilities, reading appreciation, and accuracy and consistency in thinking in problematical situations. Pupils in high school classes must be stimulated to ask more questions if the benefits of questioning, as asserted in educational theory, are to be realized.

In a few studies the questioning activity of pupils has been considered as an auxiliary factor. Eisner (18) investigated the ability of teachers to estimate the intelligence and industry of their pupils and included the factor "asks questions" as one of six criteria most closely related to such estimation. Although none of his correlations was statistically significant, he found correlations between estimates of intelligence and amount of questioning activity which ranged from 0.00 to +.50. Stevens (43) studied the use of the question as a measure of teaching efficiency. She was concerned with teachers' questions but apparently counted pupils' questions separately and remarked of one class that it was "unusual because 25 of the 34 questions were asked by the pupils." Snaddon (40) investigated the use of different types of questions in teaching the industrial arts and reported a marked absence of questions asked by the pupils. He attributed this lack of pupil questioning activity to a lack of interest.

E. The Classification of Children's Questions

One of the most persistent problems connected with studies of questioning has been that of developing some sort of system whereby questions might be classified under meaningful rubrics. The ways in which this has been attempted have been varied. One of the most frequent types of classification and the easiest to make is based on the chronological order of appearance of different question words and forms in the speech of young children. Preyer (37), Trettien (48), Stern (42), Norsworthy and Whitley (33), and Kenyeres (29), in that order, have reported chronological ages at which infants were first observed to make use of certain interrogative words and forms.

Results of these studies have been reported above, but, considering the technique of classification, the impression is gained that the age of first appearance of various interrogations may perhaps be due to; (*a*) chance in observation, (*b*) the environment of the child, or (*c*) an order of desires rather than one of language or thought development. This latter possibility has been pointed out by Wallon (50). Other factors may also be involved. In most of the observations reported, data were based on the behaviour of single subjects.

Snyder (41) described four types of questions in which the interrogative word as such is last to appear. The four, in order of appearance, were: (*a*) simple variations of the imperative,—requests with a questioning intonation; (*b*) requests for permission; (*c*) requests for permission and approbation; (*d*) questions of fact, which may be subdivided into the half-question, half-exclamation types, as, "*Big noise?*", and questions involving interrogative words. Smith, in the earlier of her two studies (38), added to these four a fifth category,—requests for permission, approbation, or corroboration. Sully (45) recognized but two categories for the questions of young children, namely, (*a*) questions implying the thirst for fact, and (*b*) questions implying a desire for reason and cause.

A number of investigators have classified children's questions according to the objects about which they were asked. Hall and Smith (23), Mau (30), Downing (17), Finley (20), Palmer (34), Pollock (36), Curtis (12), Thompson (47), and Dale (13), reported investigations in which this method of classification was used. Except for the study by Hall and Smith, which dealt with curiosity as a psychological problem, these studies were primarily concerned with the building of interest-centered curricula or textboks.

Classification of children's questions on the basis of the type of thinking being done by the questioner have not been common. The most comprehensive of all systems of this sort was published by Piaget (35) in 1926, and has been employed in researches by Davis (14), McCarthy (31), Smith (39), and the writer (19). Piaget's analysis was based on 1,225 questions which a six-year old boy asked a nurse while on two-hour daily walks over a period of 10 months. At the beginning of the observation only "why" questions were recorded, later all questions were included except those of the inflected form and the assertions with an interrogative suffix; as, "Can't I?" or "Doesn't it?" Consequently the "why" group is disproportionately large and comparable percentages not meaningful.

Questions were first divided into two groups: "why's" and questions other than "why." The distinction depends simply on the presence or absence of the word "why".

"Why" questions were then subdivided further into "why's" of causal explanation, "why's" of human motivation, "why's" of justification, and "why's" of logical justification. Questions other than "why's" were divided into several groups and sub-groups which included: questions of causal explanation; questions of reality and history, including those of facts and events, place, time, modality, and invention and imagination; questions about human actions and intentions, questions about rules and usage, and questions of classification and calculation.

"Why's" of causal explanation and questions of causal explanation were distinguished from one another by the presence or absence of the word "why". The same method was used to arrive at a distinction between "why's" of human motivation and some of the questions about human actions and intentions.

An objection of Piaget's treatment of "why" questions on the grounds that the "jars and jolts," "breakdowns and confusions" which give rise to the "why" were underestimated was made by Isaacs (27). From the standpoint of a logician and without including any illustrative questions in his discussion, Isaacs suggested the following reorganized classification of "why" questions: first, "affective and expressional why's," expressing protest, vexation, helplessness, surprise, recognition, and so forth; second, "epistemic why's," including questions of epistemic puzzlement and explanation forcing some manner of reorientation, readjustment, or rationalization; third, "informational why's," and fourth, "justificatory why's."

The difference between the system of classification for "why" questions which Isaacs proposed and that developed by Piaget appears to lie chiefly in the point of view from which questions are regarded. Piaget's classification was made in terms of the anticipated answers to the questions; Isaacs, on the other hand, thought in terms of the psychological state of the questioner. Piaget may have slighted somewhat the confusion aspect of "why" questions but he did not overlook it.

Since the classification proposed by Piaget was reported another system has been developed by Gatto (21) from an analysis of the classroom questions of several hundred elementary school pupils. The system contains 23 rubrics designed to detect the study activities

suggested by pupils' questions. The categories included are: (1) memory for a specific fact or facts from material already studied, (2) memory involving the organization of facts having a quality of unity, (3) memory for specific facts and causal reasoning (a fact and the reason for the fact), (4) memory for organization of fact, evaluation, and causal reasoning, (5) memory for organization of fact, evaluation, and inference, (6) memory for specific fact or facts not based directly on material already studied, (7) recall of information leading to comparison, (8) recall of information leading to causal reasoning, (9) information and causal reasoning (both the fact and the reason are asked for), (10) causal reasoning from material not previously studied, (11) comparison, (12) organization of fact (imaginative) and inference, (13) organization of fact (imaginative) and causal reasoning, (14) criticism or evaluation, (15) criticism or evaluation and comparison, (16) criticism or evaluation, inference and causal reasoning, (17) analysis, (18) interpretation of meaning, (19) picture study identification, (20) picture study, description of content, (21) picture study, interpretation of activity, (22) map study, organization of fact, (23) map study, location of specific points. By a grouping of these categories, Gatto arrived at eleven specific types of study activities implied by the questions. Additional cross-referencing to detect duplications might have enhanced the functional value of the system.

Summary

The term "question" has been variously defined but the meaning of the word has commonly been assumed without attempts at definition. The common elements of such definitions as have appeared include: (*a*) an element of confusion, (*b*) an awareness of a lack of specific information, and (*c*) a realization of gaps in relationships. "Question" and "problem" have not been adequately differentiated, although there seems to be a more complex connotation to the meaning of "problem."

Most of the early writers explained children's questioning activity in terms of an "instinct of curiosity." Later writers have avoided reference to any "instinctive" origins of questioning activity and have also shown less inclination to regard children's questions as implying other states than curiosity. The latest assertion as to the origin of questioning is that questions are illustrative of curiosity which is "an activity attitude very much like interest."

An abundance of theoretical assertions as to the merits of questioning activity have appeared. Such assertions usually designate the question as: (*a*) functional in the thinking process, (*b*) preparing the questioner for the response, (*c*) inhibiting opposing interests or activities, (*d*) motivating purposeful activities, (*e*) diagnostic of learning difficulties, and (*f*) a means of extending knowledge. The claim is commonly made that the conventionalized school routine suppresses the "natural" tendency to question frequently. Nearly all writers agree that children in school should be stimulated to ask more questions and many feel that in the interest of more rapid learning, teacher questions could often be replaced by pupil questions.

The earliest and most numerous statistical analyses of questioning activity were concerned with the chronological ages at which various question forms and interrogative words were first observed among infants learning to talk. There was no great agreement among reporters as to the order of appearance of the different interrogatives except that "why" questions appeared later than questions involving the other common interrogatives.

Carefully controlled investigations of the questioning activity of pupils in classrooms or children of school age have been rare. Most of those reported have dealt with studies of children's interests, commonly with the purpose of designing interest-centered curricula and textbooks. Some writers have described significant differences in the amount and type of questioning activity which can be explained on the basis of sex, age, intelligence, interests, measures of thinking, and socio-economic environment. One writer has related classroom questions to study activities. Another has shown evidence that children achieve better school learning when encouraged to ask questions in the classroom.

None of the classification systems which have been designed to separate pupils' questions into distinctive types are completely defensible from logical or practical viewpoints. Each has had advantages for its own particular purpose.

G. Some Needed Research on Children's Questioning Activity

Many problems related to children's questioning activity have been inadequately answered or entirely ignored. Much additional research must be completed before the full implications of such activity can be known. Some suggested questions for further investigation are:

1. How may the function of the question in the processes of thought and learning be isolated and its educational implications be ascertained?

2. What are the relative pedagogical merits of questions arising spontaneously among pupils in contradistinction to skillfully adapted questions asked by the teacher and accepted as problems for solution by the pupils?

3. What is the possibility of training pupils to ask a greater number of intelligent questions and what effect has such training on the development of the ability to think more clearly?[3]

4. To what extent is the amount and type of questioning activity dependent on non-intellectual personality or character traits?

5. How validly may spontaneous pupil questions be used to indicate interest or critical thinking?

6. Is there a definite developmental hierarchy in child thought of various forms of interrogatives and what relationship may exist between such a hierarchy and mental maturation?

7. How do individuals compare or vary in the extent and characteristics of their questioning activity as they move from one educational setting (teacher, classroom, class group, subject-matter area, and the like) to another and what relationships do such comparisons or variations bear to other factors.

8. May the questions of pupils be validly treated as entities reflecting original context even when removed from it and, if so, what constitutes a meaningful classification by which types of questions may be distinguished?

There are undoubtedly other aspects of the general problem of children's questioning activity which are neither touched in the literature or named above but which are deserving of investigations. The eight problems just described appear to the writer as of major significance, and any reliable investigation of them should be a valuable contribution to educational literature, theory, and technique.

References

1. Bain, A. Education as a Science. New York: D. Appleton, 1879.
2. Berry, P. G. A practical course in household chemistry. *J. Educ. Res.* 1922, **5**, 125-34.

[3]Some observations on this problem were made by Helseth (25), but her data were gathered from a very small population.

3. BOHANNON, E. W. Peculiar and exceptional children. *Ped. Sem.*, 1896, **4**, 3-60.

4. BOHN, WM. B. A child's questions. *Ped. Sem.*, 1916, **23**, 120-22.

5. BOYD, WM. Development of sentence structure in childhood. *Brit. J. Psychol.*, 1926-27, **17**, 181-191.

6. BRANDENBURG, G. C. The language of a three-year-old child. *Ped. Sem.*, 1915, **22**, 89-120; 1919, **26**, 27-40.

7. BRIGGS, THOS. H. Naïve questions and science teaching. *Sch. Sci. & Math.*, 1923, **23**, 34-40.

8. CLAPAREDE, E. La psychologie de l'intelligence. *Scienta*, 1917, 161-3.

9. COHEN, F. S. What is a question? *Monist*, 1929, 39, 350-64.

10. COMPAYRE, G. Development of the Child in Later Infancy. New York: Appleton, 1902.

11. COREY, S. M., & FAHEY, G. L. Inferring mental complexity from pupils' classroom questions. *J. Educ. Psychol.* (In press.)

12. CURTIS, F. D. Some values derived from extensive reading in general science. *Contrib. Educ., Teach. Coll., Columbia Univ.*, 1924, No. 163.

13. DALE, E. Children's questions as a source of curriculum material. *Educ. Res. Bull.*, 1937, **16**, 57-66.

14. DAVIS, E. A. The form and function of children's questions. *Child Devel.*, 1932, **3**, 57-74.

15. DEWEY, J. Democracy and Education. New York: Macmillan, 1916.

16. DOUGLASS, H. Modern Methods in High School Teaching. New York: Houghton Mifflin, 1926.

17. DOWNING, E. R. Children's interests in nature materials. *Nat. Stud. Rev.*, 1912, **8**, 334-38.

18. EISNER, H. The classroom teacher's estimation of intelligence and industry of high school students. *Contrib. Educ., Teach. Coll., Columbia Univ.*, 1937, No. 726.

19. FAHEY, G. L. A study of the classroom questions of high school pupils and the relation between questioning activity and various other factors of educational significance. Unpublished Ph.D. dissertation, University of Wisconsin, 1939.

20. FINLEY, C. W. Some studies of children's interests in science materials. *Sch. Sci. & Math.*, 1921, **11**, 1-24.

21. GATTO, F. M. Pupil's questions: Their nature and their relationship to the study process. Unpublished Ph.D. dissertation, University of Pittsburgh, 1929.

22. GRIFFITH, C. R. Psychology Applied to Teaching and Learning. New York: Farrar & Rinehart, 1929.

23. HALL, G. S., & SMITH, T. L. Curiosity and interest. *Ped. Sem.*, 1903, **10**, 315-58.

24. HASSLER, J. O., & SMITH, R. E. The Teaching of Secondary Mathematics. New York: Macmillan, 1930.

25. HELSETH, I. O. Children's thinking: A study of the thinking done by a group of grade children when encouraged to ask questions about United States history. *Contrib. Educ., Teach. Coll., Columbia Univ.*, 1926, No. 209.

26. HOLLINGSWORTH, H. L. Mental Growth and Decline. New York: D. Appleton, 1928.

27. ISAACS, SUSAN. Intellectual Growth in Young Children with an Appendix on Children's "Why" Questions by Nathan Isaacs. New York: Harcourt, Brace, 1930.

28. JOHNSON, H. M. Children in the Nursery School. New York: John Day, 1928.

29. KENYERES, E. Mots de l'enfant et l'apparition des especes de mots dans son language. *Arch. de Psychol.*, 1927, **20**, 191-218.

30. MAU, L. E. Some experiments regarding the relative interests of children in the physical and biological materials in the kindergarten and primary grades. *Nat. Stud. Rev.*, 1912, **8**, 285-291.

31. McCARTHY, D. Language development of the preschool child. *Mono-Series No. IV*, Institute of Child Welfare, University of Minnesota, Minneapolis, Minnesota, 1930.

32. NICE, M. M. Concerning all day conversations. *Ped. Sem.*, 1920, **27**, 166-77.

33. NORSWORTHY, N., & WHITLEY, M. T. Psychology of Childhood. New York: Macmillan, 1926.

34. PALMER, E. L. How to meet some of children's nature interests. *Nat. Stud. Rev.*, 1922, **23**, 23-30.

35. PIAGET, J. The Language and Thought of the Child. New York Harcourt, Brace, 1932.

36. POLLOCK, C. A. Children's questions as a basis of what to teach in general science. *Educ. Res. Bull.*, 1924, **3**, 3-6.

37. PREYER, W. The Development of the Intellect. New York: D. Appleton, 1889.

38. SMITH, M. E. Investigation of the development of the sentence and extent of vocabulary in young children. *University of Iowa Studies in Child Welfare*, 1926.

39. ———. The influence of age, sex, and situation on the frequency, form, and function of questions asked by preschool children. *Child Devel.*, 1933, **4**, 201-13.

40. SNADOON, G. H. How to use the question in teaching the industrial arts. *Indust. Arts Mag.*, 1927, **16**, 451-3.

41. SNYDER, A. D. Notes on the talk of a two-and-one-half-year old boy. *Ped. Sem.*, 1914, **21**, 412-24.

42. STERN, W. Psychology of Early Childhood. New York: Holt, 1924.

43. STEVENS, R. The question as a measure of efficiency in instruction. *Contrib. Educ., Teach. Coll., Columbia Univ.*, 1912, No. 48.

44. STORM, H. C. Who can ask the better questions—teachers or pupils? *Elem. School J.*, 1928, **28**, 610-15.

45. SULLY, J. Studies of Childhood. New York: D. Appleton, 1896.

46. SYMONDS, P. E. *Education and the Psychology of Thinking.* New York: McGraw-Hill, 1936.

47. THOMPSON, L. H. Children's questions. *Educ. Res. Bull.*, 1924, **3**, 347.

48. TRETTIEN, A. W. Language interests of childhood. *Ped. Sem.*, 1904, **2**, 113-77.

49. UMSTATTD, J. G. Secondary School Teaching. Boston: Ginn, 1937.

50. WALLON, H. L'interrogation chez l'enfant. *J. de Psychol.*, 1924, **21**, 170-82.

51. WASHBURNE, C. W. Scientific Method in the Construction of School Textbooks. New York: World Book, 1920.

52. WESLEY, E. B. Teaching the Social Studies. New York: Heath, 1937.

53. WOODWORTH, R. S. Dynamic Psychology. New York: Columbia Univ. Press, 1922.

54. WRINKLE, W. L., & ARMENTROUT, W. D. Directed Observation and Teaching in Secondary Schools. New York: Macmillan, 1932.

55. YOAKAM, G. A., & SIMPSON, R. G. An Introduction to Teaching and Learning. New York: Macmillan, 1934.

The University of Pittsburgh
Johnstown Center
Johnstown, Pennsylvania

The Journal of Genetic Psychology, 1942, 60, 359-373.

A STUDY OF STANFORD-BINET VOCABULARY ATTAINMENT AND GROWTH IN CHILDREN IN THE CITY OF CHILDHOOD, MOOSEHEART, ILLINOIS, AS COMPARED WITH CHILDREN LIVING IN THEIR OWN HOMES*

The Mooseheart Laboratory for Child Research, and the Merrill-Palmer School

VIRGINIA VAN DYNE FLEMING[1]

While the writer was working as a research assistant at the Mooseheart Laboratory for Child Research, she was impressed with the fact that the vocabulary scores on the Stanford-Binet Test (1916 form) seemed in many cases to be lower than one would have predicted from a knowledge of the chronological and mental ages. To test this impression a study of Mooseheart cases was made, and a year later, when it was possible to get parallel data on children living in their own homes, a comparative study of the two groups was made. The Mooseheart cases include all those in the Mooseheart files at the time of collecting data which satisfied the criteria set up for the study; the Home cases, similarly, include all those in the files of the Merrill-Palmer School which met these criteria.

A. ENVIRONMENT OF CHILDREN STUDIED

1. *Mooseheart Cases*

In general, the children at Mooseheart come from homes that are inferior to the environment offered by this "City of Childhood," as is shown in Reymert and Hinton's (1) study, which also points out that the average economic level of the fathers of his random Mooseheart cases is low. This study gives the mean educational level of the mothers of this group as Grade 8. Mooseheart is a

*Received in the Editorial Office on January 17, 1941.

[1]The sincere thanks of the writer are given to Dr. Martin L. Reymert, Director of the Mooseheart Laboratory for Child Research, for permission to pursue there the study here reported and to use the laboratory data for that purpose, also for his thorough criticism of the manuscript. The writer wishes to thank also Dr. Katherine E. Roberts, Chairman of the Research Committee of the Merrill-Palmer School, for her untiring criticism and help in writing this study, and Dr. Russell C. Smart of the Merrill-Palmer School, who was the able statistical adviser.

home and school for dependent children of deceased members of the Loyal Order of Moose. No crippled or feeble-minded children are admitted, and all the children of a family who are less than 14 years old are admitted with their mothers. Children at Mooseheart range in age from infancy to 18½ years, when they are graduated. According to the Reymert-Hinton study, the *IQ* distribution at Mooseheart is the same as is found in any normal community. The average *IQ* of the Mooseheart cases used in this study lies within the average range.

At Mooseheart the physical supervision, recreation, and vocational training are obviously superior to those which prevail in most of the homes from which the children come. The schools are annually approved by the proper accrediting agencies. In studying personality Thomson (4) found that the average personality quotient for his 259 Mooseheart subjects was 94, in contrast with 100 for children in their own homes, and concludes that "there is no statistically significant difference which would indicate that the Mooseheart environment is more restrictive in nature than the average home environment."

2. *Merrill-Palmer Cases*

In general, these children come from high average or superior homes, as judged by socio-economic status. Some of them attend private schools; most attend public schools. These children have had exceptional opportunities. In addition to the cultural background, physical supervision, and educational advantages of these children, they have also attended the Merrill-Palmer Nursery School, and most of them were attending a Merrill-Palmer Recreational Club at the time the mental tests used in this study were made. Their average *IQ* was "superior."

Throughout the study it must be kept in mind that the Merrill-Palmer children, referred to as "Home children," were living in their own homes, and that the Mooseheart children were not.

B. Choice of Subjects

The children studied were classified in six groups, as follows:

1. *Mooseheart Groups*

a. Mooseheart Admission. This group consisted of 46 children from Mooseheart, none of whom was over 14 years old at the time

the admission Stanford-Binet test was given (within less than a year after admission, or, more usually, a little before entrance into Mooseheart), and had a score of at least 20 words in the vocabulary portion of that test, using both word lists, or twice the score made on the one word list administered (more usually the former). The eldest child in each family was selected in recording vocabulary growth, and siblings of that child were excluded from the study.

b. Mooseheart Residence. This group consisted of 148 Mooseheart children, each of whom had had a Stanford-Binet test sometime after entering the "City of Childhood" (at least a year after), at which time he was not over 14 years of age and had scored at least 20 words in the vocabulary test. No one was included in this group who had been included in the Admission group, or who had a sibling in that group.

c. Mooseheart Growth. This group was made up of those children in Groups 1 and 2 who had been given two or more Stanford-Binet tests, at least a year apart, for each of which the child was not over 14 years of age and had passed at least 20 words in the vocabulary test. Family connections were disregarded. For each child the vocabulary gain was figured between each two tests given. In the event that one child had had more than two of these tests, his record yielded more than one usable "vocabulary growth"; therefore, one child's record could contribute more than one of the 87 cases which finally comprised this group.

2. *Home Groups*

a. Home Admission. This group consisted of 54 children, each of whom had had a Stanford-Binet test a little before, or within less than a year after, entrance into the Merrill-Palmer Recreational Clubs, at which time he was not over 14 years of age, and had a score of at least 20 words in the vocabulary portion of that test. The eldest child in each family was selected, and siblings of that child were excluded.

b. Home Residence. This group consisted of 43 children in the Merrill-Palmer Recreational Clubs, each of whom had had a Stanford-Binet test sometime after entering the Club (at least a year after), at which time he was not over 14 years of age and had a score of at least 20 words in the vocabulary test. No one was included in this group who had been included in the Home Admission group, or who had a sibling in that group.

c. Home Growth. This group was made up of 78 children in Groups 4 and 5 who had been given two or more Stanford-Binet tests, at least a year apart, for each of which the child was not over 14 years old and had passed at least 20 words in the vocabulary test. Family connections were disregarded. Vocabulary gain was figured as in the Mooseheart Growth group.

In the Home groups, most of the children had had several Stanford-Binet tests before reaching the age at which they scored at least 20 words on the vocabulary test. Therefore, in the Home Admission group, the tests used are the first ones which meet the chronological age and vocabulary criteria (under 14 years, at least 20 words), regardless of how many tests previously had been given. Hence, "entrance" into the Recreational Club does not mean his actual entrance into the Club, as it does entrance into Mooseheart, but rather the time at which the first usable test was given during his membership in the Club.

Throughout the study the usual abbreviations of *CA* for chronological age, *MA* for mental age, and *IQ* for intelligence quotient, are used. The vocabulary or raw score is the number of words passed in the two lists of the Stanford-Binet vocabulary test, or twice the number passed in the one list given. The interpolated vocabulary age is arrived at by the assumption that where two years elapse between the norms given in the test for vocabulary score at any two given age levels, the difference in norm is evenly distributed throughout that two-year interval. Thus: where raw score is: 20, 25, 30, 35, 40, 45, 50, 58, 65, 70, 75 words interpolated vocabulary age is: 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 years.

C. Results

1. *Correlations of Raw Score and Mental Age*

Table 1 gives the correlation of raw score and mental age for the four groups studied for this point. Terman (3) found that, in a group of 631 pupils, from first grade to freshman in high school, ranging in *IQ* from less than 50 to more than 150, the vocabulary score correlated with *MA* .91. In the present study, the range of *IQ*'s is smaller. While all the correlations are less than .91, only those for the Mooseheart and Home Admission groups are significantly lower, and that may be partly a function of the shorter *IQ* range.

TABLE 1

Group	N	r	$\pm$
Mooseheart Admission	46	.72	.051
Mooseheart Residence	148	.88	.014
Home Admission	54	.64	.055
Home Residence	43	.86	.027

2. *Judging Mental Age from Interpolated Vocabulary Age*

Table 2 is a frequency table showing the amount and direction of all errors of judgment of mental age when the latter is based

TABLE 2

ERRORS IN JUDGING MENTAL AGE FROM INTERPOLATED VOCABULARY AGE IN FOUR OF THE GROUPS STUDIED, WITH SEXES LISTED SEPARATELY

Error (years)	Moose. Adm. Boys	Moose. Adm. Girls	Moose. Res. Boys	Moose. Res. Girls	Home Adm. Boys	Home Adm. Girls	Home Res. Boys	Home Res. Girls
—4.6 to —5.0			1					
—4.1 to —4.5		1						
—3.6 to —4.0								1
—3.1 to —3.5	1		1	1				3
—2.6 to —3.0			3	5	1	1	1	1
—2.1 to —2.5	1	1	4	8	2	3	3	2
—1.6 to —2.0	3	2	9	5	1	1	3	2
—1.1 to —1.5	3	2	11	9	5	4	7	3
— .6 to —1.0	4	1	14	11	5	4	1	2
— .5 to — .1	6	4	20	10	7	2	2	0
0	1	0	3	2	0	0	0	0
.1 to .5	4	1	14	1	5	2	3	3
.6 to 1.0	4	2	6	2	3	2	3	2
1.1 to 1.5	4		4		3	1		
1.6 to 2.0	1		3		1			
2.1 to 2.5				1				1
2.6 to 3.0					1			

solely upon a knowledge of interpolated vocabulary age. The errors are stated in terms of years. Where the interpolated vocabulary age would lead us to believe that the child is mentally younger than he is, the figures in the "error" column bear a minus sign; where it would lead us to believe that he is older, no sign accompanies the figures. In other words, a minus indicates a vocabulary that is deficient for the mental age; no sign indicates a vocabulary that is superior for the mental age.

An analysis of Table 2 yields the following information regarding the percentage of cases in each group in which the interpolated

vocabulary age equals, exceeds, or is exceeded by the mental age (Table 3). In more than half the cases in all four groups, the

TABLE 3

Moose. Adm:	Moose. Res:	Home Adm:	Home Res:	
63%	77%	67%	72%	interpolated vocabulary age is exceeded by *MA*
3	3	0	5	interpolated vocabulary age equals *MA*
35	20	33	23	interpolated vocabulary age exceeds *MA*

interpolated vocabulary age alone would indicate a mental age lower than the actual one. Considering mental age, vocabulary would often seem to be deficient in all four groups.

Terman (2) writes that the Stanford-Binet vocabulary test alone can give the mental age correctly within a year in about 60 per cent of the cases; and within a year and a half in 80 per cent of them. These figures are based upon the previously mentioned study by Terman (3) in which the *IQ* range was greater than in the present study, and in which he used a different scale than the one used here to estimate the interpolated vocabulary age ("mental age" in his scale). Terman's corresponding scale is:

Median vocabulary:	13	18	23	30	35	41	46	51	57	62	67	73	75
Mental age:	7	8	9	10	11	12	13	14	15	16	17	18	19

Scrutiny of the material in this study reveals the following data regarding this subject (Table 4). It would seem that, on the whole,

TABLE 4

	Moose. Adm:	Moose. Res:	Home Adm:	Home Res:	Terman:
Correct within ½ year	35%	34%	30%	19%	
" " 1 "	59	56	56	37	about 60%
" " 1½ year	78	73	80	61	" 80
" " 2 "	91	84	85	72	
" " 2½ "	95	92	94	86	
" " 3 "	96	99	100	91	
" " 3½ "	98	100	100	98	
" " 4 "	98	100	100	100	
" " 4½ "	100	100	100	100	
" " 5 "	100	100	100	100	

Terman's estimates of prediction accuracy are essentially valid for these groups of children, especially when the differences in *IQ* range

and in scale of interpolated vocabulary age are considered. They are least true of the Home Residence group.

These figures also indicate that the range of errors in prediction is somewhat less for the Home than in the Mooseheart groups. Errors in judging mental age of the Home cases range from +3.0 to —4.0; those in the Mooseheart cases range from +2.5 to —5.0. Thus, while the range of errors of judgment of mental age from vocabulary ability is about the same in both Home and Mooseheart cases, the figures indicate a tendency toward greater vocabulary deficiency in the Mooseheart cases.

3. *Sex Differences*

Since the data could easily be analyzed for purposes of studying quantitative sex differences in vocabulary, and since many people are interested in such differences in relation to factors connected with quantitative vocabulary differences, the following analysis is included, although it was not initially a part of the problem specifically studied in this piece of work.

The mean and sigma for *CA, MA, IQ,* and vocabulary score were computed for the Mooseheart and Home, Admission and Residence groups, with sexes kept separate. These data are found in Table 5. The sex differences in vocabulary score in each of the four groups are statistically unreliable. Terman (3) also found sex differences so slight as to have little or no significance.

TABLE 5

MEAN *CA, MA, IQ,* AND VOCABULARY SCORE, FOR MOOSEHEART AND HOME, ADMISSION AND RESIDENCE GROUPS, WITH SEXES KEPT SEPARATE

Sex	*N.*	Years residence at Moose. M	σ	*C.A.* *M*	σ	*M.A.* *M*	σ	*IQ* *M*	σ	Vocabulary score *M*	σ
						Mooseheart Admission					
Boys	32			12.2	.4	11.4	2.5	94	13.7	20	4.7
Girls	14			11.9	1.4	11.4	2.0	96	14.0	16	5.1
						Mooseheart Residence					
Boys	93	7.8	3.3	12.4	1.4	11.5	1.8	92	13.0	18	10.1
Girls	55	7.3	2.8	12.1	1.6	11.5	2.1	95	14.0	17	8.8
						Home Admission					
Boys	34			8.0	1.3	9.8	1.3	122.4	12.2	30.6	6.1
Girls	20			8.5	1.7	10.2	2.1	115.6	12.8	30	7.4
						Home Residence					
Boys	23			11.6	2.1	14.0	2.1	120.7	15.0	47.2	11.8
Girls	20			10.9	1.7	13.5	2.6	125.5	19.1	45	11.8

TABLE 6

ACCURACY OF JUDGMENT OF *MA* FROM INTERPOLATED VOCABULARY AGE ALONE, IN MOOSEHEART AND HOME, ADMISSION AND RESIDENCE GROUPS, SEXES SEPARATE

Judgment correct within	Mooseheart Adm.		Mooseheart Res.		Home Adm.		Home Adm.	
	32 boys %	14 girls %	93 boys %	55 girls %	34 boys %	20 girls %	23 boys %	20 girls %
½ yr.	34	36	40	24	35	20	22	15
1 yr.	59	57	62	47	59	50	39	35
1½ yrs.	81	71	78	64	82	75	70	50
2 yrs.	94	86	91	73	88	80	83	60
2½ yrs.	97	93	95	89	94	95	96	75
3 yrs.	97	93	98	98	100	100	100	80
3½ yrs.	100	93	99	100				95
4 yrs.		93	99					
4½ yrs.		100	99					
5 yrs.			100					

Table 6 shows the accuracy of judgments of mental age made solely from interpolated vocabulary age, in the four groups, with sexes figured separately.

In general, prediction seems to be more accurate in Home than in Mooseheart cases. Judging from the data in Table 3, it would seem that sex differences as to accuracy of prediction are greater in Mooseheart than in Home cases, especially in the Residence group.

So much for the accuracy of prediction in terms of the amount of error made. As for the direction of that error, the material below is given. For, although we may know that a judgment is correct within two years, it is important to know also whether our judgment gives the child a mental age which is two years too little, or too much, judging from vocabulary ability. In these data, a minus sign indicates a vocabulary deficiency. Judging by mental age, a plus sign indicates a better vocabulary than mental age would lead us to expect, and a zero indicates that the two are in exact agreement (Table 7). In all but one of the four groups (Home Residence) the deficiency is greater for girls than for boys. The results in Home groups closely resemble those in corresponding Mooseheart groups, the greatest difference being between girls in the Mooseheart and Home Residence groups. The difference in mean *IQ* is also greatest between these two groups.

TABLE 7

	Moose. Adm. Boys	Moose. Adm. Girls	Moose. Res. Boys	Moose. Res. Girls	Home Adm. Boys	Home Adm. Girls	Home Res. Boys	Home Res. Girls
—	56	79	68	89	62	75	74	70
0	3	0	3	4	0	0	0	0
+	41	21	29	7	38	25	26	30

4. *Quantitative Vocabulary Resemblances in Siblings*

All siblings used in any of the Mooseheart groups were listed, and the discrepancy between interpolated vocabulary age and mental age was figured for each Stanford-Binet score. Where a child had had more than one test, the average discrepancy was used. In each pair of siblings the difference between them in this discrepancy was computed. Where there were three or more siblings, the difference between the two most discrepant averages was used. The mean difference in the amount of discrepancy between interpolated vocabulary age for the Mooseheart families was 1.16 years. This difference was considered small enough to exclude siblings in the parts of the study where siblings were included. The corresponding figure for the Home cases was 1.12 years.

5. *Effect of Length of Stay in Mooseheart upon Vocabulary Score*

In Table 8 is presented material which shows the relation of varying lengths of residence in Mooseheart to vocabulary growth, using the 148 Mooseheart Residence cases. The vocabulary score figures in Table 8 are an average of all the cases which fell within the designated portion of the table. Mental age and years of residence are in whole numbers, so that a figure 7, for example, includes from 7 to 7.9. It appears from this table that within the span of one year's mental age the benefit in vocabulary growth received from Mooseheart training does not vary in any regular or proportional amount with the length of time spent there. Given two children of the same general intelligence, as measured by mental age, we are quite unable to predict their vocabulary scores according to that mental age or to show any consistent effect of a known number of years of residence in Mooseheart.

6. *Vocabulary Growth in Mooseheart and Home Cases*

In the Mooseheart and Home Growth groups the gain in vocabulary score between each two Stanford-Binet tests was figured. Many

TABLE 8

AVERAGE VOCABULARY SCORES OF THE 148 MOOSEHEART RESIDENCE CASES, GROUPED ACCORDING TO MENTAL AGE AND NUMBER OF YEARS IN RESIDENCE AT MOOSEHEART

Mental age	Number of years in residence 1	2	3	4	5	6	7	8	9	10	11	12	13
7		10											
8			10	10	10	10.5	14						
9	10	12.5	10	16	12	11	11.5	13	12.5	15		13.5	
10	18.5	14	17	16	14	14.5	16	16	13.5	14.5			
11	19		19	21.5	20.5		17.5	26	17	19	17.5	17.5	18
12			20	15	19	18	21		19	23	24	20.5	
13			23	23		15	22.5	19.5		20	30	25	23.5
14		22.5			21	26	22					21.5	
15					25		27.5						
16				25									

of the tests were given a year apart, and in that case the gain, or loss, was easily figured. However, where more than one year had elapsed between two tests, the vocabulary growth (in terms of number of words) was divided by the number of months elapsing between the two tests, and the result was multiplied by 12, to give a probable annual growth. The 87 Mooseheart and 78 Home Growth cases were divided into three intelligence level groups on the basis of all individual and group tests administered to each case. Then the average gain in vocabulary per year was computed for each classification. These data are shown in Table 9. These data indicate that for both Mooseheart and Home children the average and below-average intelligence groups had an annual vocabulary growth rate that was significantly substandard, while the children of above-average intelligence were growing at a rate not significantly below the standard. The average annual vocabulary growth rate for Mooseheart children of average intelligence was 38 per cent of the standard; for Home children, it was 62 per cent of the standard. The average vocabulary growth rates for below-average and average Home children, and above average Mooseheart children, are practically the same. The difference found between Mooseheart and Home children of above-average intelligence is statistically significant; the Home children have a larger annual vocabulary growth rate. These findings would seem to indicate either that Terman's standards are correct, and all of the groups involved in the present study are deficient in vocabulary, or that Terman's standards in this respect are too high. In this connection it is important to note that the environment of the Home children is one which is usually thought to be exceptionally conducive to vocabulary development. This environment includes above average socio-economic level and its usual results, recreational reading and story-telling, nursery school attendance, and broad social experiences.

7. *Correlation of Vocabulary with Ability to Name Words at Random for Three Minutes*

For 162 Home cases, the correlation of the number of words named in three minutes (test No. 6, year level *X*) with mental age was .499 ± .040. For 148 Home cases, the number of words named correlated with interpolated vocabulary age .508 ± .041. (Terman, 3, found the correlation to be .49.) Attainment in this

TABLE 9

AVERAGE ANNUAL VOCABULARY GROWTH RATE FOR THREE DIFFERENT INTELLIGENCE LEVELS FOR 87 MOOSEHEART AND 78 HOME CASES

		N	Mean annual vocabulary growth	*SD* annual vocabulary growth	*SE* of mean	*S-M* difference between mean and standard	*SM* critical ratio for significance of difference between mean and standard	*D*, Mooseheart-Home	*SE* difference	Critical ratio for significance between Mooseheart and Home
Below average	Mooseheart	16	1.0	1.2	.29	4.0	13.8	2.0	2.49	.68*
	Home	2	3.0	3.5	2.47	2.0	.8			
Average	Mooseheart	62	1.9	3.2	.41	3.1	7.6	1.2	.05	1.35
	Home	10	3.1	2.5	.79	1.9	2.4			
Above average	Mooseheart	9	3.1	2.9	.97	1.9	2.0	2.6	1.08	2.41
	Home	66	5.7	3.8	.47	.7	1.5			

*Using Fisher's formula for critical ratio in small samples

test seems as closely correlated to mental age as to interpolated vocabulary age.

D. Summary

The Stanford-Binet tests (1916 form) for 194 Mooseheart and 97 Merrill-Palmer, or Home, children were analyzed in an effort to compare vocabulary growth. Various correlations and growth norms were compared with those given by Terman. Vocabulary score as a predictor of mental age was studied. Sex differences in the various groups were analyzed. Selection of cases to be studied involved getting an estimate of quantitative vocabulary resemblances among siblings. The relationship between vocabulary growth and length of stay at Mooseheart was studied. A comparison was made between vocabulary growth in 87 Mooseheart and 78 Home cases. In addition to the points included in the conclusions, the following points cover the findings of this study:

1. The vocabulary of these children as they entered Mooseheart was less closely related to *MA* ($r=.72$) than Terman's correlation of .91 might indicate. This was true also of the children as they entered the Merrill-Palmer Recreational Clubs ($r=.64$). For those children who had been in Mooseheart, or in the Clubs, at least a year, vocabulary was related to *MA* to about the degree indicated by Terman (.88 and .86, respectively).

2. Of the children studied upon entrance into Mooseheart, 63 per cent had vocabularies which were deficient for their *MA*. Of the children who had been in Mooseheart more than a year, 77 per cent had deficient vocabularies (according to Stanford-Binet test requirements). The vocabularies of the Home children were deficient to a similar degree.

3. In general, for both the Mooseheart and the Home cases, vocabulary was essentially as accurate a tool for predicting *MA* as Terman indicates.

4. The range of error in predicting *MA* from vocabulary attainment was somewhat less in the Home than in the Mooseheart cases. The direction, rather than the amount, of the error of prediction indicated a tendency toward greater vocabulary deficiency in the Mooseheart than in the Home cases.

5. Siblings appeared to have quantitatively similar vocabulary-*MA* differences. This sibling similarity was practically identical in the Mooseheart and Home cases (1.16 and 1.12 years, respectively).

6. Judgments of *MA* from vocabulary score were almost as accurate for children who had been in Mooseheart for a year or more as for the children as they entered Mooseheart.

7. The mean *IQ* for children entering Mooseheart was 94 for the boys, 96 for the girls; for children who had been there for a year or more, 92 for the boys, 95 for the girls; for children entering the Merrill-Palmer Recreational Clubs, 122 for the boys, 116 for the girls; for children who had been in these Clubs at least a year, 121 for the boys, 126 for the girls.

8. In the Institutional and Home, Admission and Residence cases, there were no reliable sex differences in vocabulary scores.

9. In the Institutional Admission and Residence groups, and in the Home Admission group, vocabulary age was lower than *MA* in a higher percentage of the girls than of the boys. In the Home Residence group these figures showed scarcely any sex difference.

10. The percentage of children who were deficient in vocabulary, considering *MA*, was greater among those who had been in Mooseheart a year or more than it was among those who had just entered. To a lesser degree the same was true of the children entering the Merrill-Palmer Clubs and those who had been there at least a year.

11. In comparing Mooseheart and Home children, practically the same percentage of cases had a vocabulary which was better than their *MA* would have led us to predict.

E. Conclusions

1. For both Mooseheart and Home children, only the above-average intelligence group had an annual vocabulary growth rate not differing significantly from the standard.

2. The average and below-average intelligence groups of both Mooseheart and Home children were significantly substandard in their annual vocabulary growth rate.

3. In all three intelligence groups the average annual growth rate in vocabulary was greater for Home than for Mooseheart children. However, this difference was statistically significant only in the above-average group.

4. The rate of vocabulary growth for below-average and average Home children and above-average Mooseheart children was practically identical.

5. In general, however, in Mooseheart the higher the intelligence

level the greater was the vocabulary growth per year. This was true also of the Home children.

6. Comparison of children of similar mental age indicated that the benefit in vocabulary growth received from Mooseheart training did not vary in any regular or proportional amount with the length of time spent in the City of Childhood.

References

1. Reymert, M. L., & Hinton, R. T., Jr. The effect of a change to a relatively superior environment upon the IQ's of one hundred children. *Intelligence: Its Nature and Nurture.* 39th Yearbook, Nat. Soc. Stud. Educ., 1940, Part II, 255-268.
2. Terman, L. M. The Intelligence of School Children. New York: Houghton Mifflin, 1919. Pp. 317.
3. ———. Vocabulary test as a measure of intelligence. *J. Educ. Psychol.*, 1918, 9, 452-467.
4. Thomson, W. An evaluation of the *PQ* (Personality Quotient) Test. *Charac. & Personal.*, 1938, 4, 274-292.

The Merrill-Palmer School
71 East Ferry Avenue
Detroit, Michigan

The Journal of Genetic Psychology, 1942, 60, 375-380.

THE SOCIO-ECONOMIC STATUS OF THE HOMES OF MENTALLY SUPERIOR AND RETARDED CHILDREN AND THE OCCUPATIONAL RANK OF THEIR PARENTS*

North Carolina State College of A. and E., and the Murray State Teachers College

WILLIAM McGEHEE AND W. D. LEWIS[1]

The investigations, with the notable exception of that of Stoke and Lehman (7), which have dealt with the problem of the relationship of children's intelligence to the socio-economic status and occupational rank of parents have been so preoccupied with the positive relationships found that they have failed to emphasize adequately the many exceptions reported in their data. As Stoke and Lehman (7) have indicated, in the subsequent reporting of the results of these investigations, there has been a tendency to omit the exceptions and to build up the concept that perfect or almost perfect correlations exists between occupational rank and intelligence or between socio-economic status and intelligence. Secondhand accounts find quotable passages such as the one from Jordan (3, p. 117) that "median scores (on intelligence tests) increase *pari passu* with economic level" and overlook entirely the implications of the word *median*.

Further, the generality of data on the relationship of intelligence to social status as secured in the majority of investigations is open to question either from the point of view of the technique used in securing the subjects, as in the Terman (8) study, or due to the relatively limited geographic area, if not number of cases, from which subjects are drawn, as in Haggerty and Nash's (2) investigation.

It is believed that this investigation presents evidence which underscores the position taken in the introductory paragraph that investigators should emphasize negative as well as affirmative evidence of relationship between intelligence and social status. It is

*Received in the Editorial Office on February 8, 1941.

[1]The writers acknowledge their indebtedness to the advisory committee of the Coördinated Studies in Education, Inc., for permission to use the data presented in this investigation.

believed, further, that the data to be presented are less open to the criticism of prejudicial selection of subjects, than previous investigations have been, as the subjects are drawn from a sample of approximately 45,000 children in grades four through eight in 455 schools, in 310 communities, and in 36 states in the United States.

As part of the Coördinated Studies in Education,[2] the teachers of the children involved in the study were asked to indicate on previously prepared blanks the father's present or last regular occupation and the material aspects of the child's home. The items included under the material aspects of the home were possession and number of telephones, automobiles, radios, regular servants; whether or not the family subscribed to a newspaper; the number of rooms in the house and the number of people occupying each room. The present investigation will compare children in respect to the just cited factors who, on the basis of the Kulhmann-Anderson Tests scores, can be designated as mentally superior or mentally retarded.

Mentally superior children in this investigation are those children who, on the basis of the Kulhmann-Anderson Tests, scored in the upper 10 per cent of the entire subject population of approximately 45,000 children. Data on occupational rank of parents are available for 4,176 children in this group and on socio-economic status, for 4,237 subjects. Mentally retarded children in this study are those who scored in the lowest 10 per cent of the parent population on the basis of scores on the Kulhmann-Anderson Tests. Data on occupational status are available for 3,697 children in this group and, on socio-economic status, for 3,571 subjects. Data on parental occupation and socio-economic status also are presented on a third group of children taken at random from the larger subject population and assumed to represent the "average" child in the study. Data on parental occupation of 12,390 children in this group and on the socio-economic status of 12,018 children are presented. Evidence is available that this sample of "average" children, over one-fourth of the total population, is characteristic of the entire population and that no significant differences are to be expected if the entire parent population had been used as the "average" group.[3]

[2]The nature and extent of the Coördinated Studies in Education has been described in detail previously and therefore will not be repeated. See (4) and (5) in bibliography.

[3]It has not been considered necessary to give the separate data on sex and grade groups, as the differences are all small and statistically unreliable. The data on both parental occupations and socio-economic status have been treated, therefore, for both sexes and all grades combined.

TABLE 1

THE INTELLIGENCE QUOTIENTS IN TERMS OF THE 90TH, 50TH, AND 10TH PERCENTILE OF THE CHILDREN IN GRADES FOUR THROUGH EIGHT

%iles	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
90	116.8	115.5	116.6	118.1	115.0
50	101.8	100.7	101.6	101.6	99.3
10	82.2	81.0	82.8	86.1	86.7
N	10,230	8,939	10,187	8,193	7,944

Table 1 shows the lower limits of the superior group, the upper limits of the retarded group, and the median of the entire population in terms of *IQ*'s.

The father's occupations in the present investigation have been rated on the basis of the Terman-Taussig classification (8). The five categories in this classification are: I. Professional; II. Semi-professional and business; III. Skilled labor; IV. Semi-skilled and slightly skilled labor; and V. Common labor. The classification of the occupational levels of the fathers of the subjects in the three groups in this investigation is shown in Table 2.

TABLE 2

A COMPARISON OF THE OCCUPATIONAL LEVELS OF THE FATHERS OF THE SUPERIOR, AVERAGE, AND RETARDED GROUPS, GIVING PERCENTAGES: ALL GRADES AND BOTH SEXES COMBINED

Group	Occupational levels					No. of cases
	I	II	III	IV	V	
Superior	10.3	35.9	37.4	10.3	5.9	4,176
Average	4.3	22.1	42.6	11.2	19.5	12,390
Retarded	.6	10.9	41.6	15.6	28.9	3,697

It is evident from the data presented in Table 2 that a positive relationship does exist between occupational status of a child's parent and the child's intelligence. This relationship is indicated even more clearly in Table 3. In this table are presented the ratios of children in the superior and retarded groups at each occupational level in terms of normal expectancy as indicated by the "average" group.

The fact which should be emphasized from the data in Table 2 and Table 3, however, is that superior and retarded children are found at all occupational levels. Further, at Level III, which

TABLE 3

THE RATIO OF SUPERIOR AND RETARDED CHILDREN AT EACH OCCUPATIONAL LEVEL IN TERMS OF NORMAL EXPECTANCY BASED ON DATA FROM AVERAGE CHILDREN: ALL GRADES COMBINED

Group	Occupational level I	II	III	IV	V
Superior	2.40	1.62	.88	.92	.30
Retarded	.14	.49	.98	1.39	1.53

contributes over one-third of the children to all three groups, the differences between percentage of children in each of the groups are small and the difference is not statistically significant between the average and retarded group. It is to be further noted that this group, furnishing the largest percentage of both superior and retarded children has only 88 per cent of normative expectancy in the case of superior children and 98 per cent in the case of retarded subjects. Groups II, III, and IV, "the great middle class," furnish approximately 84 per cent of all the superior subjects and approximately two-thirds of the retarded group. It may be true, as Collins (1, p. 169) contends, that "the occupations of the father, then, may be considered as a rough index of the intelligence of the child." The index, however, is so "rough" as to have practically no diagnostic significance.

The data on fathers' occupation and material aspects of the home have been combined to secure a socio-economic index for the subjects in this investigation. The combined data have been rated in terms of a socio-economic rating schedule which is a modification of the Sims *Score Card for Socio-Economic Status* (6). The following credits were given for the various occupational levels of fathers:

I.	Professional	4
II.	Business	3
III.	Skilled labor	2
IV.	Semi-skilled labor	1
V.	Common labor	0

One point each was given if the home had a telephone, automobile, radio, regular servant, or newspaper; two points were given for two or more of any one of these. The room-per-person ratio was given the points indicated in Table A.

A comparison of the median socio-economic rank of the three

TABLE A

Ratio	Points
.5 or less	0
.51 - 1.0	1
1.01 - 1.5	2
1.51 - 2.0	3
2.01 or more	4

TABLE 4

A COMPARISON OF THE SOCIO-ECONOMIC RATINGS OF THE SUPERIOR, AVERAGE, AND RETARDED GROUPS, GIVING MEDIANS AND QUARTILES: ALL GRADES AND BOTH SEXES COMBINED

Group	Md.	Q_3	Q_1	N
Superior	8.09	10.05	6.04	4,237
Average	6.61	8.50	4.87	12,018
Retarded	4.18	6.05	2.61	3,571

groups in this study (Table 4) shows what is usually found when the central tendencies of intelligence test scores of different socio-economic groups are compared. The superior children have a median socio-economic score higher than average children, and average children, a median socio-economic score higher than retarded children. The differences between the median socio-economic scores of these groups are statistically reliable. Taken on the face-value of *median* scores, higher intelligence is positively associated with socio-economic status.

It is valuable, however, to consider the distribution of the children at the various socio-economic ranks before over-emphasizing the importance of the differences in these medians. Table 5 presents these data.

The inference from the data in this table is showing that superior

TABLE 5

A COMPARISON OF THE PERCENTAGE OF HOMES OF THE SUPERIOR, AVERAGE, AND RETARDED RATED AT DIFFERENT SOCIO-ECONOMIC LEVELS

Type of home	Socio-economic score	Superior group	Average group	Retarded group
Poor	0-3	8.0	15.0	30.7
Average	4-8	53.4	65.0	61.8
Average-Superior	7-11	55.0	55.0	22.9
Superior	12-18	5.9	2.0	0.7

and retarded children come from all types of homes. They come from the poorest as well as the best, but the great bulk of them come from "average" homes. The middle-class homes rating from 4-8 inclusive, furnish 65 per cent of the school population and 85 per cent of the normative expectancy of superior children, 96 per cent of the normative expectancy of retarded children. The upper middle class homes, rating from 7 to 11 inclusive, furnish 55 per cent of the school population and 100 per cent normative expectancy of superior children, 46 per cent normative expectancy of retarded children. It is evident, therefore, that atypical children are found in all types of homes and that the relationship between socio-economic status of the home and the intelligence of the child living in the home is by no means high.

In conclusion, then, this study has indicated that positive relationships between socio-economic status and mental ability holds primarily when group averages are considered. It has indicated that a knowledge of the parent's occupation or the socio-economic condition of the home of the child is a very precarious index of the child's intelligence. It would even suggest that investigators of the Iowa persuasion should exercise extreme caution in attributing any reported increase in test intelligence solely to environmental changes.

References

1. Collins, J. E. Intelligence of school children and parental occupation. *J. Educ. Res.*, 1928, **17**, 157-169.
2. Haggerty, M. E., & Nash, H. B. Mental capacity of children and parental occupations. *J. Educ. Psychol.* 1924, **15**, 559-572.
3. Jordan, A. M. Parental occupations and children's intelligence scores. *J. Appl. Psychol.*, 1933, **4**, 353-366.
4. Lewis, W. D. A study of superior children. *George Peabody Coll. Contrib. Educ.*, 1940, No. 266. Pp. 104.
5. McGehee, W. A study of retarded children. *George Peabody Coll. Contrib. Educ.*, 1939, No. 246. Pp. 128.
6. Sims, V. M. The Measurement of Socio-Economic Status. Bloomington, Ill.: Public School Publishing, 1928. Pp. iii+33.
7. Stoke, S. M., & Lehman, H. C. Intelligence test scores of social and occupational groups. *Sch. & Soc.*, 1930, **31**, 372-377.
8. Terman, L. M. Genetic Studies of Genius. Vol. I. Stanford University: Stanford Univ. Press, 1925. Pp. vii+648.

Department of Psychology
North Carolina State College of A. and E.
Raleigh, North Carolina

Department of Education
Murray State Teachers College
Murray, Kentucky

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1942, **60**, 381-384.

A STUDENT EVALUATION OF COURSE OBJECTIVES IN PSYCHOLOGY*

Stephens College

CHARLES J. MARSH[1]

Most instructors have a good idea of what their aims or objectives are in the teaching of psychology. Nevertheless, it is difficult to state these objectives in simple language. An attempt was made to do this by the psychology staff of Stephens College during a recent Faculty Conference week. After considerable labor, a set of 15 objectives was drawn up, more as a guide for future effort than as a statement of what was already being accomplished.

At the end of a course in General Psychology, 150 students were asked to indicate the extent to which each of these objectives had been met in her particular case. These 150 students were freshman and sophomore women, taking their first course in Psychology at Stephens College. None of them had seen the set of objectives previously.

The purpose of this paper is to present these 15 objectives for inspection and comment by other teachers of General Psychology, and to present student opinion upon the success with which these objectives were judged to have been met. The directions to the students were:

> You are asked to give your candid opinion of the degree to which these objectives have been met. Please give each item your critical thought. An honest opinion is essential. Anything else would be misleading, and therefore worse than worthless.
>
> Whatever you say here will not have the slightest effect on

*Received in the Editorial Office on February 5, 1941.

[1]Grateful acknowledgment is made to Drs. Constance Chandler, C. N. Rexroad, and J. C. Stauffacher for their contributions in setting up the objectives and in gathering data on student reactions.

your status—either in the eyes of the instructor or in the eyes of the administration. You may sign the blank if you wish, but this is not required.

At the left of each objective are five numbers: 1, 2, 3, 4, 5. After reading each objective, consider to what degree it has been achieved *for you*, and then draw a circle around one of the numbers to indicate your judgment. Use the following scheme:

1 means that the objective has been met in a *completely* satisfactory way.
2 means that the objective has been *very well* met.
3 means that the objective has been *fairly well* met.
4 means that the objective has been *only slightly* achieved.
5 means that the objective has been *missed entirely*.

Table 1 shows the list of objectives, the total number of students

TABLE 1
AVERAGE RATINGS ON FIFTEEN OBJECTIVES

Objectives	*N*	Average rating
Development of a sympathetic understanding of the faults and peculiarities of other people. An objective attempt to understand the real causes of the faults of others, and a considerate acceptance of such persons rather than prejudiced behavior toward them.	149	1.7
A genuine effort toward self-improvement psychologically.	148	1.9
Reasonably good and reasonably rapid adaptation to changes in environment.	147	2.0
Recognition of your own prejudices and a genuine effort to minimize their undesirable effects.	149	2.0
Emotional control. Progress toward the understanding of emotional causes and effects and toward the elimination of merely impulsive action.	149	2.1
Assumption of the responsibility for one's own actions, decisions and limitations.	147	2.1
Reaction to others on the basis of the whole personality rather than on the basis of a single prominent trait.	147	2.1
Formation of sound judgments. This includes a critical attitude toward rumor and gossip, an insistence on command of facts before forming an opinion, and restraint in forming strong likes and dislikes.	148	2.2
Recognition of the prejudices of others, especially the ability to recognize and discount "stereotypes," and the ability to evaluate statements of others on the basis of their soundness rather than on the prestige of the person who makes the statement.	148	2.2

TABLE 1—Continued

Objectives	N	Average rating
A development of interest in psychological issues shown in reading and thinking beyond the required assignments. This would include news items and magazine articles which were appropriate, or it might include popular or technical books, or psychological novels.	150	2.2
Valid self-understanding. This would involve discovery of genuine assets and liabilities; comprehension of the true basis of the attitudes of others; development of valid self-confidence and personal adjustment.	149	2.4
Friendly interest in people without imposing your own personality upon them.	145	2.4
Differentiation between real explanation and mere description, and the ability to discriminate sound from faulty generalization.	148	2.4
Intelligent planning. This would include an objective analysis of abilities and interests; consideration of remote as well as immediate consequences of decisions; reliable estimates of the relative values of various courses of action.	148	2.4
A development of interest in psychological issues shown in pertinent contributions to discussion groups and in attendance at lectures or study groups.	100	2.5

who responded to each objective, and the average rating received by each objective. The table is arranged as a series, with the objective which received the highest average rating at the top and the objective which received the lowest average rating at the bottom. It is to be noted that this arrangement does not imply that those objectives which appear near the top of the table are considered more important than the others, either by the staff or by the students. In the present study, there is no information available on the relative importance of the 15 objectives. When a similar study is made with a new group of students, it is planned to have the importance of the objectives rated, as well as the degree to which each is considered to have been achieved.

The average rating for all the objectives taken together was 2.2, which can be interpreted as meaning that the average student believes that the 15 objectives set up for the Psychology course have been "very well met." The mean rating of 2.2 is very close to this category.

Further analysis of the obtained data was made by totalling up all the ratings for each category, and converting to percentages. Table 2 shows these results.

TABLE 2
PERCENTAGES OF JUDGMENTS IN EACH CATEGORY FOR ALL OBJECTIVES TAKEN TOGETHER

Rating:	1	2	3	4	5	Total
Percentage:	24.8%	40.0%	26.4%	6.6%	2.2%	100.0%

Table 2 indicates that approximately one-quarter of the students considered the objectives to be met in a "completely satisfactory way" and approximately one-quarter considered them to be "fairly well met." About 40 per cent were of the opinion that the objectives were "very well met" while less than 10 per cent considered the objectives to be "only slightly" achieved or "missed entirely."

In general those objectives dealing with the subject of personality and emotion were judged as most successfully met, and those objectives dealing with less personal issues were judged as least successfully met. This finding is consistent with the emphases of the course.

SUMMARY

One hundred fifty college women who had just completed their first course in General Psychology were asked to express their opinions as to how well a set of 15 objectives had been met for them. Results indicate that the group felt that the objectives had been "very well met" in general, and that those objectives dealing with personality and emotions had been more successfully met than some other objectives.

Stephens College
Columbia, Missouri

The Journal of Genetic Psychology, 1942, 60, 385-386.

A NOTE ON THE WRENN STUDY HABITS INVENTORY*

Educational Records Bureau, New York City

ARTHUR E. TRAXLER

The Wrenn *Study Habits Inventory* is a diagnostic device for use in analyzing the good and the poor study habits of high-school and college students. It consists of 30 statements to which the students respond by checking (*a*) rarely or never, (*b*) sometimes, or (*c*) often or always. Eight of the statements are grouped under the general heading of note-taking and reading techniques, 10 statements are concerned with habits of concentration and school interests, and 12 statements pertain to general habits of work.

The responses are scored with a weighted key and thus a total study habits score may be obtained for each student. The relationship of the scores thus obtained to other measures is a matter of some interest for counselors, remedial reading teachers, and others using the inventory.

The Wrenn *Study Habits Inventory;* the Otis *Self-Administering Test of Mental Ability, Higher Examination;* the *Iowa Silent Reading Test, Advanced Form A;* the Whipple *High School and College Reading Test;* the Booker *Reading Test;* and the Inglis *Test of English Vocabulary* were administered to 76 university freshmen. The median *IQ* of the group was 116.5. The correlation between the scores on the Wrenn *Study Habits Inventory* and the scores on the other tests are shown in Table 1.

The correlations for study habits and intelligence and study habits and vocabulary are slightly negative. The correlations between the study habits inventory and the reading tests are positive but low. None of the correlation coefficients is statistically significant.

These correlations indicate that whatever is measured by the Wrenn *Study Habits Inventory* is not related to any significant degree to intelligence, reading comprehension, reading rate, or English vocabulary. The lack of significant relationship between the study habits scores and the reading scores is somewhat surprising. Although

*Received in the Editorial Office on February 20, 1941.

TABLE 1

Correlation of Wrenn Study Habits Inventory With Certain Intelligence, Reading, and Vocabulary Tests

Test	Wrenn Study Habits r PE
Otis *IQ*	—.144 ± .076
Iowa Reading Comprehension	.139 ± .076
Iowa Reading Rate	.159 ± .075
Whipple Reading	.130 ± .076
Booker Reading Comprehension	.067 ± .077
Booker Reading Rate	.227 ± .073
Inglis Vocabulary	—.008 ± .077

these data cannot be regarded as evidence concerning the validity of the inventory for the purposes intended, they do suggest the desirability of cautious interpretation of the inventory scores in connection with a remedial reading program.

It is obvious, of course, that the responses to the individual items in the inventory may have important diagnostic and guidance values aside from the usefulness of the scores on the inventory as a whole.

Educational Records Bureau
437 West 59 Street
New York City

The Journal of Genetic Psychology, 1942, 60, 387-394.

BOOKS

The *Journal of Genetic Psychology*, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than $2 per printed page and not more tran $3 per printed page.

Conditions. Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (*a*) Makes an important theoretical contribution; (*b*) consists largely of original experimental research; (*c*) has a creative or revolutionary influence in some special field or the entire field of psychology; (*d*) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. Just send in your review. It does not matter if the book in question has been reviewed before.

(1940)

CANTRIL, H. The Invasion from Mars: A Study in the Psychology of Panic. Princeton: Princeton Univ. Press, 1940. Pp. 228.

RASHEVSKY, N. Advances and Applications of Mathematical Biology. Chicago: Univ. Chicago Press, 1940. Pp. 214.

BURLINGAME, L. L. Heredity and Social Problems. New York: McGraw-Hill, 1940. Pp. 369.

DENNIS, W. The Hopi Child. New York: D. Appleton-Century, 1940. Pp. 204.

HILGARD, E. R., & MARQUIS, D. G. Conditioning and Learning. New York: Appleton-Century, 1940. Pp. 429.

HARTMAN, G. W., NEWCOMB, T., *Editors*. Industrial Conflict: A Psychological Interpretation. New York: Cordon, 1940. Pp. 583.

DOOB, L. W. The Plans of Men. New Haven: Yale Univ. Press, 1940. Pp. 411.

HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. B. Mathematico-Deductive Theory of Rôte Learning. New Haven: Yale Univ. Press, 1940. Pp. 329.

LANDIS, C., *et al.* Sex in Development. New York: Hoeber, 1940. Pp. 329.

THORNDIKE, E. L. Human Nature and the Social Order. New York: Macmillan, 1940. Pp. 1019.

GESELL, A., *et al.* The First Five Years of Life. New York: Harper, 1940. Pp. 393.

YOUNG, K. Personality and Problems of Adjustment. New York: Crofts. Pp. 868.

BROWN, J. F. The Psychodynamics of Abnormal Behavior. New York: McGraw-Hill, 1940. Pp. 484.
GOLDSTEIN, K. Human Nature. Cambridge: Harv. Univ. Press, 1940. Pp. 258.
KATONA, G. Organizing and Memorizing. New York: Columbia Univ. Press, 1940. Pp. 318.
BAKKE, E. W. The Unemployed Worker. New Haven: Yale Univ. Press, 1940. Pp. 465.

(1941)

ZIPF, G. K. National Unity and Disunity. Bloomington, Ind.: Principia Press, 1941. Pp. 408.
REIK, T. Masochism in Modern Man. New York: Farrar & Rinehart, 1941. Pp. 439.
BRITT, S. H. Social Psychology of Modern Life. New York: Farrar & Rinehart, 1941. Pp. 562.
BARTLEY, S. H. Vision: A Study of Its Basis. New York: Van Nostrand, 1941. Pp. 350.
CANTRIL, H. The Psychology of Social Movements. New York: Wiley, 1941. Pp. 274.
MILLER, N. E., & DOLLARD, J. Social Learning and Imitation. New Haven: Yale Univ. Press, 1941. Pp. 341.
DEARBORN, W. F., & ROTHNEY, J. W. M. Predicting the Child's Development. Cambridge: Sci-Art 1941. Pp. 360.
FARAGO, L. (Ed.) German Psychological Warfare. New York: Committee for National Morale, 1941. Pp. 133.

(1942)

SHELDON, W. H. The Varieties of Temperament. New York: Harper, 1942. Pp. 520.
KLOPFER, B., & KELLEY, D. M. The Rorschach Technique. New York: World Book, 1942. Pp. 436.
SCHILDER, P. Goals and Desires of Man. New York: Columbia Univ. Press, 1942. Pp. 305.
MILLER, J. G. Unconsciousness. New York: Wiley, 1942. Pp. 329.
ALLPORT, G. W. The Use of Personal Documents in Psychological Science. New York: Social Science Research Council, 1942. Pp. 210.
BENDER, I. E., IMUS, H. A., ROTHNEY, J. W. M., KEMPLE, C., & ENGLAND, M. R. Motivation and Visual Factors. Hanover, N. H.: Dartmouth College, 1942. Pp. 369.
BORING, E. G. Sensation and Perception in the History of Experimental Psychology. New York: Appleton-Century, 1942. Pp. 644.

CRITICAL REVIEWS OF RECENT BOOKS

(Dennis, W. The Hopi Child. New York: Appleton-Century, 1940. Pp. 204.)

REVIEWED BY HELEN WOLFLE

The specific purpose of this review is to emphasize the opportunities given a Hopi infant for learning to walk. It is not a general review of *The Hopi Child,* which has aspects to interest the nutritionist concerned with infant welfare, the anthropologist concerned with the breakdown of Indian under the impact of American customs, and other features. Rather, it is a review of the specific data of most interest to the psychologist—the data concerning walking and the use of the cradleboard.

The Hopi parent's purpose in using the cradleboard is not to provide a maturation experiment for American psychologists, but to provide a suitable environment for Hopi babies. The cradleboard "is to ensure that the child will be straight and of good carriage" (p. 31), but "after three months the straightness of the child is assured and the mother may discard the cradleboard whenever she pleases" (p. 33). Besides the interest in good posture, the Hopi watches the development of walking and notes its proper timing. Since the purpose of using the cradleboard is fulfilled earlier, it need not interfere with the later development of walking.

The child is bound to the cradle during the first day. First it is wrapped in a blanket much as any newborn baby is, and then strips of cloth are tied around both child and cradle (p. 31). "The ties are passed over the legs also, so that they can be flexed only slightly" (p. 31). "When the infant is tied to the cradleboard on the first day of life the process does not cause the baby to cry" (p. 95).

From birth to three months, the infant is almost always on the cradleboard. It "is employed very assiduously for the first three months, the child being taken off the board only for cleaning and bathing, acts which combined do not occupy more than an hour daily" (p. 33). When the infant "is kept off of the board for a longer period than is usual he cries and does not go to sleep" (p. 96).

From three months to six months, "the child enjoys freedom of movement for a larger and larger part of the day" (p. 97).

> "After three months . . . the mother may discard the cradleboard whenever she pleases. Actually, its use is seldom discontinued before the child is six months of age and it rarely is employed beyond the first year of life. The duration of the cradle usage depends in part upon the restlessness of the child and in part upon his motor development. If a child becomes restless on the board, he is freed earlier than would otherwise be the case" (p. 33).

During this period the baby begins a life outside the house without the cradleboard.

> "The child is seldom taken from his home until he is four months of age . . . He is ordinarily carried in arms or upon the back of the mother, and the cradleboard is left at home" (p. 96). "An older sister of the child, if there be one, takes over his care to a greater and greater degree. She watches over him, carries him about, and takes him out of the house for a part of the day. In carrying the baby, the older child places him on her back and then puts around him a shawl, the ends of which she holds securely in her hands. The baby should not be carried until he is three or four months of age. Before this age, he is not considered strong enough" (p. 35).

In Chicago, the approved practice for the airing of a four-months-old baby is to wrap him in a zippered case without sleeves, cover him with a varying number of blankets depending on the weather, and place him on the back porch in a buggy. There he is left for his morning nap. By comparison, the jaunts of the Hopi baby must involve considerably more fatiguing exercise.

From six months to nine months, the Hopi baby is given still greater freedom. "After six months of age he is often placed on his back on the floor in order that he may exercise himself" (p. 35). "The infant must not be placed on his abdomen before he can get himself into this position, for to do so will hurt the child. The child must not be tossed in the air in play until he is eight or ten months of age" (p. 34). "Babies are petted and fondled, they are placed on the lap, and are carried about" (p. 34) at this and other ages.

From nine months to fourteen months, the baby learns to walk. "When he is old enough, he is led by the hand to encourage walking. He is expected to walk shortly after he is one year old. If he should be late in starting to walk, his legs may be rubbed early in the morning and he may be bathed in cold water" (p. 35).

Further evidence of the Hopi's interest in having the child learn to walk is found in the accuracy with which the date of first walking is remembered. "When asked at what age a given child began to walk, the Pueblo mother, after a few moments, gave a reply such as, 'A year and one month and a few days,' or 'Two days before the first birthday' or something equally circumstantial" (p. 106).

It has usually been discarded earlier, but the cradleboard *must* be discarded after the child walks. "One mother ceased using the cradleboard when her son was fourteen months of age. For a week he slept badly and cried at night. Another mother whom we knew also discontinued the use of the cradleboard when her son was fourteen months of age" (p. 96). Now and then the attachment to the cradleboard is so strong that the infant is allowed to sleep on it after this age. "It must not be thought that all infants are as fond of the cradleboard as those to whom we have just referred" (p. 97).

The restraint imposed by the cradleboard is summarized in these words: "The cradling of the Hopi infant prevents a number of actions which our babies commonly engage in, such as bringing the hands to the mouth, playing with the hands, watching the hands, and putting the feet in the air" (pp. 98-99). Even so, "observation of a large number of Hopi infants of various ages shows that the behavior which is characteristic of the different ages is approximately the same" for American infants.

> "In spite of the enforced extension of the limbs, the Hopi infant when freed from his bindings for the bath or for the changing of bedding, takes the usual flexed position. Although his hands are held downward perhaps twenty-three hours in twenty-four, when he is at liberty he puts them to the mouth and carries objects to his mouth as do white babies. He reaches for objects and handles them at approximately the same time as do white children. He reaches for his toes and puts his toes in his mouth. Sitting, creeping, and walking follow the usual sequence" (p. 101).

As is well known, another group of Hopi infants were not kept on the cradleboard. Their average age of walking was almost identical with the average age of infants using the cradling device. Since the differences are quite insignificant, there is no evidence of an effect of the cradleboard upon walking" (p. 107).

These are interesting data, and the reviewer objects to none of the careful conclusions made about them. The title of an earlier

article, *The effect of cradling practices upon the onset of walking in Hopi children* (1) gives an invitation to the hasty reader to suppose that the cradleboard should or might have an effect. If one escapes this error, one may suppose, as Ruch (2) does, that since walking was not delayed by the cradleboard, walking is therefore primarily a function of maturation and is relatively independent of practice. This conclusion finds no support in the data. The proper conclusion to make is that the cradleboard, as used by the Hopi, does not significantly interfere with the practice of those coördinations which lead to walking.

It has been shown that the Hopi concerns himself with the walking of his child, encouraging the process in several ways and carefully noting the day of its occurrence. "From three months of age onward the child enjoys freedom of movement for a larger and larger part of the day" (p. 97). "Sitting, creeping and walking follow the usual sequence" (p. 101). It is difficult to imagine how the cradleboard could have made any difference in age of walking, and the facts indicate that it does not.

References

1. Dennis, W., & Dennis, M. G. The effect of cradling practices upon the onset of walking in Hopi children. *J. Genet. Psychol.*, 1940, **56**, 77-87.
2. Ruch, F. L. Psychology and life. (Rev. ed.) Chicago: Scott, Foresman, 1941. Pp. 754.

5225 Drexel Avenue
Chicago, Illinois

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i.e., the most recently received books will be found at the end of the list.)

ROETHLISBERGER, F. J. Management and Morale. Cambridge: Harvard Univ. Press, 1941. Pp. 194.

OPLER, M. E. An Apache Life-Way. Chicago: Univ. Chicago Press, 1941. Pp. 500.

ALEXANDER, F. M. The Universal Constant in Livnig. New York: Dutton, 1941. Pp. 270.

HART, H. New Gateways to Creative Living. New York: Abingdon, 1941. Pp. 197.

ROBINSON, F. P. Diagnostic and Remedial Techniques for Effective Study. New York: Harper, 1941. Pp. 318.

WANG, C. K. A. An Annotated Bibliography of Mental Tests and Scales. Peiping: Catholic Univ. Press, 1940. Pp. 698.

MORGAN, J. J. B. Workbook in Psychology. New York: Farrar & Rinehart, 1941. Pp. 122.

FALK, H. A. Corporal Punishment. New York: Teach. Coll., Columbia Unlv., 1941. Pp. 162.

TERMAN, L. M., & HOWARD, L. S. The Road Ahead. New York: World Book, 1941. Pp. 402.

HAYES, S. P. Contributions to a Psychology of Blindness. New York: Amer. Found. for the Blind, 1941. Pp. 296.

LENDE, H. Books About the Blind. New York: Amer. Found. for the Blind, 1941. Pp. 215.

HARRIS, R. M. Teachers' Social Knowledge and Its Relation to Pupils' Responses. New York: Teach. Coll., Columbia Univ., 1941. Pp. 89.

CHATTERTON, R. H. Methods of Lesson Observing by Preservice Student-Teachers. New York: Teach. Coll., Columbia Univ., 1941. Pp. 137.

SPADINO, E. J. Writing and Laterality Characteristics of Stuttering Children. New York: Teach. Coll., Columbia Univ., 1941. Pp. 82.

WILLMOTT, J. N. High School Boys Electing Industrial Arts. New York: Teach. Coll., Columbia Univ., 1941. Pp. 71.

TODD, J. E. Social Norms and the Behavior of College Students. New York: Teach. Coll., Columbia Univ., 1941. Pp. 190.

GOLDFARB, W. An Investigation of Reaction Time in Older Adults. New York: Teach. Coll., Columbia Univ., 1941. Pp. 76.

FARAGO, L. (*Ed.*) German Psychological Warfare. New York: Committee for National Morale, 1941. Pp. 133.

LEVY, RUTH J. Reductions in Recidivism through Therapy. New York: Seltzer, 1941. Pp. 143.

WHITING, J. W. M. Becoming a Kwoma: Teaching and Learning in a New Guinea Tribe. New Haven: Yale Univ. Press, 1941. Pp. 226.

PACE, C. R. They Went to College. Minneapolis: Univ. Minnesota Press, 1941. Pp. 148.

PORTERFIELD, A. L. Creative Factors in Scientific Research. Durham: Duke Univ. Press, 1941. Pp. 282.

RANK, O. Beyond Psychology. Camden, N. J.: Haddon, 1941. Pp. 291.

GROVES, C. Get More out of Life. New York: Association Press, 1941. Pp. 136.

HAYAKAWA, S. I. Language in Action. New York: Harcourt, Brace, 1941. Pp. 245.

Walton, A. Fundamentals of Industrial Psychology. New York: McGraw-Hill, 1941. Pp. 231.

Voronoff, S. From Cretin to Genius. New York: Alliance Book, 1941. Pp. 281.

Ford, C. S. Smoke from Their Fires. New Haven: Yale Univ. Press, 1941. Pp. 248.

Dearborn, W. F., & Rothney, J. W. M. Predicting the Child's Development. Cambridge: Sci-Art, 1941. Pp. 359.

Price, L. Creative Group Work on the Campus. New York: Columbia University, 1941. Pp. 437.

Reuning, K. Joy and Freude. Swarthmore, Pa.: Swarthmore College, 1941. Pp. 141.

Fitt, A. B. Seasonal Influence on Growth Function and Inheritance. Wellington: New Zealand Council for Educational Research, 1941. Pp. 182.

Sen Gupta, N. N. Heredity in Mental Traits. Calcutta: Macmillan, 1941 Pp. 207.

Angyal, A. Foundations for a Science of Personality. New York: Commonwealth Fund, 1941. Pp. 398.

Waring, E. B., & Johnson, M. W. Helping Children Learn. Ithaca; Cornell Univ. Press, 1941. Pp. 460.

Ullmann, E. V. Diet in Sinus Infections and Colds. New York: Macmillan, 1941. Pp. 185.

Hayes, S. P. Contributions to a Psychology of Blindness. New York: American Foundation for the Blind, 1941. Pp. 296.

Strong, D. S. Organized Anti-Semitism in America. Washington, D. C.: American Council on Public Affairs, 1941. Pp. 191.

Sutherland, R. L. Color, Class, and Personality. Washington, D. C.: American Council on Education, 1942. Pp. 135.

Hildreth, G. The Child Mind in Evolution. New York: King's Crown Press, 1941. Pp. 163.

Rope, F. T. Opinion Conflict and School Support. New York: Columbia University, 1941. Pp. 164.

Graebner, I., & Britt, S. H. Jews in a Gentile World. New York: Macmillan, 1942. Pp. 436.

Shaffer, L. F., Gilmer, B. von H., & Porter, J. M. Experiments and Demonstrations in Psychology. New York: Harper, 1942. Pp. 230.

American Youth Commission. Youth and the Future. Washington, D. C.: American Council on Education, 1942. Pp. 296.

Hamilton, P. Hangover Square. New York: Random House, 1942, Pp. 308.

Lundberg, G. A. Social Research. New York: Longmans, 1942. Pp. 426.

Krout, M. H. Introduction to Social Psychology. New York: Harper, 1942. Pp. 823.

Hathaway, S. R. Physiological Psychology. New York: D. Appleton Century, 1942. Pp. 335.

Sorenson, H., & Lemon, A. C. Workbook for Psychology in Education. New York: McGraw-Hill, 1942. Pp. 178.

Bowman, H. A. Marriage for Moderns. New York: McGraw-Hill, 1942. Pp. 493.

Contributions in Honor of L. M. Terman. Studies in Personality. New York: McGraw-Hill, 1942. Pp. 333.